Estimation of the Lévy Measure for the Aggregate Claims Process in Risk Theory

Md. Sharif Ullah Mozumder

A Thesis

for

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science at

Concordia University

Montréal, Québec, Canada

August 2007

Your file Votre référence
ISBN: 978-0-494-34713-3
Our file Notre référence
ISBN: 978-0-494-34713-3

ABSTRACT

Estimation of the Lévy Measure for the

Aggregate Claims Process in Risk Theory

Md. Sharif Ullah Mozumder

Lévy processes (LP) are gaining popularity in actuarial and financial modeling. The Lévy measure is a key factor in the versatility of LP applications. The estimation of the Lévy measure from data is shown to be useful in analyzing the aggregate claims processes in Risk Theory.

Starting with infinitely divisible distributions (IDD), some nice constructions are obtained for finite sums of Lévy processes. The Lévy properties of compound Poisson processes are extensively used in the thesis. Examples illustrate the close relationship between IDDs and Lévy processes.

The Poisson random measure associated with jumps of a Lévy process exceeding a given threshold is discussed and a new derivation is obtained. The relation with subordinators (increasing Lévy processes) is explored. Intuitive ideas and results are obtained for the jump function G appearing in Lévy's characterization of the Lévy–Khinchine formula. A non–parametric estimator of G is discussed. A detailed relation between G and ν, the Lévy measure, is derived, yielding an estimator of ν. The latter gives an estimator of the Poisson rate λ_ϵ and the claim size distribution F^ϵ for claims larger than the threshold ϵ. Extensive numerical simulations illustrate the paths of gamma, inverse Gaussian and α–stable claim subordinators and their corresponding estimates for λ_ϵ and F^ϵ.

Acknowledgments

This thesis was written under the joint supervision of Prof. José Garrido and Dr. Xiaowen Zhou. I am indebted to them for their cordial help. Special thanks are due to Prof. Garrido for introducing me to the nice structure of Lévy processes through this problem. I appreciate the time that Dr. Gaillardetz has given me to address many technical problems.

I remember my Mom's encouragement through all my uncomfortable times. The help and support extended by my friends Iván, Jun and Ya Fang at different times is also truly appreciated.

The financial support by Concordia University and my supervisors is highly appreciated, it made this study possible.

Thanks also go to Mrs. Anne-Marie Agnew for her many administrative and legal suggestions.

Contents

List of Figures vii

List of Tables viii

1 Tools and Characteristics 1

1.1 Introduction 1

1.2 Notion of Infinitely Divisible Distributions (IDD) 2

1.3 Examples of IDD's 5

1.4 Important Results Concerning IDD's 8

1.4.1 The Lévy measure 9

1.4.2 Lévy–Khinchine formula 10

2 Lévy Processes 15

2.1 Basic Ideas 15

2.2 Examples of Lévy Processes 19

2.2.1 The Poisson Process 19

2.2.2 Compound Poisson process 21

2.2.3 Gaussian Lévy process 24

2.3 Random Measure of Lévy Processes With Threshold Exceeding Jumps . . 25

2.4 Poisson Point Process 29

3 Special Lévy Processes and their Applications 31

3.1 Distribution of Truncated Jumps 31

3.2 Truncated Jumps Density and the Compound Poisson Process 33

3.3 Subordinators . . . 34
3.3.1 Examples of subordinators . . . 36

4 Estimation of ν and λ **41**
4.1 Insurance Aggregate Claims Process . . . 41
4.2 Some Motivations . . . 42
4.2.1 Jump function G and infinite divisibility . . . 43
4.2.2 Explanation of Property 4.1 . . . 44
4.3 Estimating G . . . 46
4.3.1 Description of the method for estimating G . . . 47
4.3.2 On the convergence of β_n . . . 53
4.3.3 Relation of G and ν . . . 54
4.4 Estimating λ_ϵ . . . 56
4.5 Claim Size Distribution Through $\hat{\nu}$ . . . 57

5 Numerical Implementation **59**
5.1 Estimation of the Rate λ_ϵ . . . 60
5.1.1 Gamma process . . . 60
5.1.2 Inverse Gaussian process . . . 60
5.1.3 α-stable process . . . 61
5.2 Approximation of Jump Size Distribution . . . 62
5.2.1 Gamma process . . . 63
5.2.2 Inverse Gaussian process . . . 64

A Preliminaries in Probability **75**
A.1 Topics from Probability Theory . . . 75
A.1.1 Notions of convergence . . . 76
A.2 Notions of Convolution . . . 80
A.3 Jumps Caglad and Cadlag Functions . . . 82
A.4 Stochastic Process . . . 84

List of Figures

3.1 Dominance of $e^{-\beta x}$ (dotted–line) over $e^{-sx}e^{-\beta x}$ for $\beta = \min\{s \mid s = 0.5$ (red), $s = 5$ (green), $s = 50$ (yellow)$\}$. 38

3.2 Dominance of the graph of $\frac{1}{x^\alpha}$ (dotted–line) over the graphs of $\frac{e^{-sx}}{x^\alpha}$ for $\alpha = \beta = \min\{s \mid s = 0.5$ (red), $s = 5$ (green), $s = 50$ (yellow)$\}$. 39

5.1 Gamma ($\alpha = \frac{0.09}{20}$, $\beta = 1$) and truncated gamma processes for the thresholds $\epsilon = 1.2, 0.5, 0.05$ and 0.01. 61

5.2 Inverse Gaussian ($\mu = 200$, $\theta = 0.3$) and truncated processes for thresholds $\epsilon = 40000, 20000, 5000$ and 100. 63

5.3 α-stable and truncated processes for thresholds $\epsilon = 1000, 500, 50, 2$. 64

5.4 Gamma ($\alpha = 1$, $\beta = 0.5$) and truncated processes for thresholds $\epsilon = 2, 1.5, 1$ and 0.5. 66

5.5 Gamma ($\alpha = 1$, $\beta = 0.5$) and truncated processes for thresholds $\epsilon = 2, 1.5, 1$ and 0.5. 67

5.6 Inverse Gaussian ($\mu = 40$, $\theta = 5$) and truncated processes for thresholds $\epsilon = 250, 200, 150$ and 100. 68

5.7 Inverse Gaussian ($\mu = 40$, $\theta = 5$) and truncated processes for thresholds $\epsilon = 250, 200, 150$ and 100. 69

5.8 Histogram of ϵ-exceeding gamma claims with estimated conditional gamma and conditional IG densities. 70

List of Tables

5.1 Estimates of rate λ_ϵ for gamma ($\alpha = \frac{0.09}{20}$, $\beta = 1$) process. 62
5.2 Estimates of rate λ_ϵ for IG ($\mu = 200$, $\theta = 0.3$) process. 62
5.3 Estimates of rate λ_ϵ for α-stable process with $\alpha = 0.8$. 65
5.4 Estimates of gamma parameters 65
5.5 Goodness of fit test for different CDF's. 65
5.6 Estimates of inverse Gaussian parameters 66
5.7 Goodness of fit test for different CDF's. 67

Chapter 1

Tools and Characteristics

1.1 Introduction

We consider a particular family of stochastic processes called Lévy processes (see Appendix A for the definition of a stochastic process on a probability space).

Definition 1.1 *A cadlag (see Appendix A) stochastic process* $X = \{X_t\,;\, t \geq 0\}$, *on* $(\Omega, \mathcal{F}, \mathbf{P})$ *with values in* $\mathbb{R}$ *is called a Lévy process if it satisfies the following properties:*

[L1] $X_0 = 0$ *a.s. (see Appendix A.1)*

[L2] X_t *has independent and stationary increments, i.e.*

(i) for every increasing sequence of times $t_0 < t_1 < t_2 < \cdots < t_n$ *the random variables* $X_{t_0}, X_{t_1} - X_{t_0}, \cdots, X_{t_n} - X_{t_{n-1}}$ *are independent.*

(ii) $X_{t+h} - X_t \stackrel{D}{=} X_{t+h-t} = X_h$ *, i.e. the distribution of* $X_{t+h} - X_t$ *does not depend on* t*.*

[L3] X_t *is stochastically continuous, i.e.*

$$\lim_{h \to 0} \mathbf{P}(|X_{t+h} - X_t| > \epsilon) = 0, \qquad \forall \epsilon > 0\,.$$

In no way does condition $[L3]$ imply that the sample paths are continuous, as we will see in the case of the Poisson process. The intuitive meaning of $[L3]$ is that for a given time

t (deterministic) the probability of seeing a jump at t is zero, i.e. discontinuities (jumps) do not occur at deterministic times and so occur at random times. It serves to exclude processes with jumps at fixed times which can be regarded as "calendar effects" and are not interesting for our modeling purposes. All these facts together with the notion of jumps (discussed in Appendix A) yield the following result.

Proposition 1.1 *If $X = \{X_t\,;\, t \geq 0\}$ is a Lévy process then for fixed $t > 0$, $\triangle X_t = 0$ a.s.*

Proof. Consider a sequence $\{t_n\,;\, n \in \mathbb{N}\}$ in $\mathbb{R}^+$ with $t_n \nearrow t$ as $n \to \infty$. Since X has cadlag paths

$$\lim_{n\to\infty} X_{t_n} = X_{t^-}.$$

However by [$L3$] the sequence $\{X_{t_n}\,;\, n \in \mathbb{N}\}$ converges in probability to X_t and so has a subsequence which converges almost surely to X_t. Hence the result follows from the definition of jumps and the uniqueness of the limits. □

Remark 1.1 *The above proposition shows that $\triangle X$ is not a straightforward process to analyze.*

We now deviate from the discussion on Lévy processes to introduce the class of infinitely divisible distributions (henceforth IDD). We will then come back to see how closely these IDD's are related with Lévy processes.

1.2 Notion of Infinitely Divisible Distributions (IDD)

The increments of a Lévy process are in one-to-one correspondence with infinitely divisible distributions. We present here a brief overview of this relation. For a more general discussion on IDD's we refer to [13], [10] and [4].

By sampling a Lévy process at times $0, \triangle, 2\triangle, 3\triangle, \cdots\cdots$ we simply obtain a random walk

$$S_n(\triangle) = \sum_{k=0}^{n-1} Y_k, \qquad \text{where each} \quad Y_k = X_{(k+1)\triangle} - X_{k\triangle},$$

are iid random variables whose distribution, by[$L2$], is the same as that of

$$Y_k = X_{(k+1)\triangle} - X_{k\triangle} \stackrel{D}{=} X_{(k+1)\triangle - k\triangle} = X_\triangle, \qquad k = 0, 1, \cdots.$$

Since this can be done for any sampling interval $\triangle$ we say that by sampling a Lévy process with different $\triangle$ we specify a whole family of random walks $S_n(\triangle)$.

Choosing $n\triangle = t$, we see that for any $t > 0$ and any $n \geq 1$,

$$\begin{aligned} S_n(\triangle) &= \sum_{k=0}^{n-1} Y_k \\ &= (X_\triangle - X_0) + (X_{2\triangle} - X_\triangle) + \cdots \\ &\qquad\qquad + (X_{(n-1)\triangle} - X_{(n-2)\triangle}) + (X_{n\triangle} - X_{(n-1)\triangle}) \\ &= X_{n\triangle} = X_t. \end{aligned}$$

That is, X_t can be represented as the sum of n iid random variables whose distribution is that of $X_\triangle = X_{t/n}$. Otherwise said, X_t is divided into n iid parts. A distribution having this property is said to be infinitely divisible. We now formally define IDD's.

Definition 1.2 *A distribution function F (or an F distributed random variable X) is said to be infinitely divisible if for any positive integer n there exists independent and identically distributed random variables $Y_1, Y_2, \cdots, Y_n$ such that $Y_1 + Y_2 + \cdots + Y_n$ is F distributed. "Equivalently" a distribution function F is infinitely divisible if and only if its characteristic function*

$$\Phi_F(s) = \int e^{isx} dF(x), \qquad s \in \mathbb{R},$$

can be written for any integer n in the form

$$\Phi_F = [\varphi]^n,$$

such that φ is also a characteristic function of some distribution.

The following result characterizes IDD's.

Theorem 1.1 *The following are equivalent:*

[1] X is infinitely divisible.

[2] F_X has a convolution n^{th} root, for any n, that itself is the distribution function of a random variable.

[3] Φ_F has an n^{th} root, for any n, that itself is the characteristic function of a random variable.

For a detailed proof we refer to [?] and [16].

Remark 1.2 In general the convolution n^{th} root of a probability measure is not unique. However it is always unique when the measure is infinitely divisible. See [7].

Now let $M_1(\mathbb{R})$ be the set of all Borel probability measures on $\mathbb{R}$. The above theorem suggests us to generalize the definition of IDD to distributions that have a convolution n^{th} root in $M_1(\mathbb{R})$.

Proposition 1.2 *$F \in M_1(\mathbb{R})$ is infinitely divisible if and only if for each $n \in \mathbb{N}$ there exists*

$$F^{1/n} \in M_1(\mathbb{R}),$$

for which

$$\Phi_F(s) = \left[\Phi_{F^{1/n}}(s)\right]^n, \qquad \text{for all } s \in \mathbb{R}.$$

Proof. Since $F \in M_1(\mathbb{R})$ is infinitely divisible there exists a convolution n^{th} root in $M_1(\mathbb{R})$ i.e.

$$F = \left(F^{(1/n)}\right)^{*n}.$$

So

$$\begin{aligned}\Phi_F(s) &= \int_{-\infty}^{\infty} e^{isx} dF(x) \\ &= \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n-times} e^{is(x_1+x_2+\cdots+x_n)} F^{1/n}(dx_1) F^{1/n}(dx_2) \cdots F^{1/n}(dx_n).\end{aligned}$$

As the random variable X with distribution function F is infinitely divisible, then in $X = X_1 + X_2 + \cdots + X_n$ each X_i has a distribution $F^{1/n}$, and it leads to

$$\Phi_F(s) = \left[\int_{-\infty}^{\infty} e^{isx_1} F^{1/n}(dx_1)\right]^n.$$

That is

$$\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n .$$

Conversely let there exist $F^{1/n} \in M_1(\mathbb{R})$ for which $\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n$. But

$$\begin{aligned}
[\Phi_{F^{1/n}}(s)]^n &= \left[\int_{-\infty}^{\infty} e^{isx} F^{1/n}(dx)\right]^n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{is(x_1+x_2+\cdots+x_n)} F^{1/n}(dx_1)\cdots F^{1/n}(dx_n) \\
&= \int_{-\infty}^{\infty} e^{isx} (F^{1/n})^{*n}(dx).
\end{aligned}$$

Comparing this expression with

$$\Phi_F(s) = \int_{-\infty}^{\infty} e^{isx} dF(x),$$

we get that $F = (F^{1/n})^{*n}$, that is F is infinitely divisible. □

1.3 Examples of IDD's

Example 1.1 Gaussian random variables

Appendix A shows that if F has the underlying random variable $X \sim N(\eta, \sigma^2)$ then

$$\Phi_F(s) = e^{is\eta - \frac{1}{2}s^2\sigma^2}, \qquad s \in \mathbb{R}.$$

So we can write

$$\Phi_F(s) = \left[e^{is\frac{\eta}{n} - \frac{1}{2}s^2\frac{\sigma^2}{n}}\right]^n, \qquad s \in \mathbb{R},$$

and hence we can recognize $F^{1/n}$ as the distribution with underlying random variable $Y \sim N(\frac{\eta}{n}, \frac{\sigma^2}{n})$ having the characteristic function

$$\Phi_{F^{\frac{1}{n}}}(s) = e^{is\frac{\eta}{n} - \frac{1}{2}s^2\frac{\sigma^2}{n}}, \qquad s \in \mathbb{R}.$$

Then $\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n$ and hence $F = (F^{1/n})^{*n}$, which implies by Proposition 1.2, that Gaussian random variables are infinitely divisible.

Similarly in the multivariate case, as shown in Appendix A, the characteristic function can be written as

$$\Phi_F(\underline{s}) = e^{i(\underline{s}.\underline{\eta}) - \frac{1}{2}(\underline{s}.A\underline{s})} = \left[e^{i(\underline{s}.\frac{\underline{\eta}}{n}) - \frac{1}{2}(\underline{s}.\frac{A}{n}\underline{s})}\right]^n, \qquad s \in \mathbb{R}^n,$$

which shows that multivariate normal random variables are infinitely divisible.

Example 1.2 Gamma random variables

If F has the underlying random variable $X \sim G(\alpha, \beta)$ then

$$\Phi_F(s) = \int_0^\infty e^{isx} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx.$$

That is

$$\Phi_F(s) = \left(\frac{\beta}{\beta - is}\right)^\alpha = \left[\left(\frac{\beta}{\beta - is}\right)^{\frac{\alpha}{n}}\right]^n, \qquad s \in \mathbb{R}.$$

Since $\left(\frac{\beta}{\beta-is}\right)^{\frac{\alpha}{n}}$ is the characteristic function of $F^{1/n} \sim Gamma(\frac{\alpha}{n}, \beta)$, we get

$$\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n \qquad \textit{implying that} \qquad F = (F^{1/n})^{*n}.$$

So gamma random variables are infinitely divisible.

Example 1.3 Poisson random variables

In the univariate case, as shown in Appendix A, if F is for an underlying random variable $X \sim Poisson(\lambda)$ its characteristic function is then

$$\Phi_F(s) = e^{\lambda(e^{is}-1)}, \qquad s \in \mathbb{R},$$

so we can write

$$\Phi_F(s) = \left[e^{\frac{\lambda}{n}(e^{is}-1)}\right]^n, \qquad s \in \mathbb{R}.$$

Thus we recognize $F^{1/n}$ as the distribution of a Poisson $(\frac{\lambda}{n})$ random variable with characteristic function $\Phi_{F^{\frac{1}{n}}}(s) = e^{\frac{\lambda}{n}(e^{is}-1)}$. Hence we get

$$\Phi_F(s) = [\Phi_{F^{1/n}}(s)]^n \qquad \textit{implying that} \qquad F = (F^{1/n})^{*n}.$$

So Poisson random variables are infinitely divisible.

Example 1.4 Compound Poisson (CP) random variables

Definition 1.3 *Suppose that $\{Z_n, n \in \mathbb{N}\}$ is a sequence of iid random variables taking values in $\mathbb{R}$ with common law F_Z and let $N \sim Poisson(\lambda)$ be independent of all Z_n. Then the compound Poisson random variable X, denoted $CP(\lambda, F_Z)$, is defined to be $X = Z_1 + Z_2 + \cdots + Z_N$, with $Z = 0$ if $N = 0$, so that we can think of X as a random walk with a random number of steps (jumps), controlled by a $Poisson(\lambda)$ random variable N and with random step sizes Z_i.*

Proposition 1.3 *For $X \sim CP(\lambda, F_Z)$ and each $s \in \mathbb{R}$*

$$\Phi_X(s) = \mathbb{E}\left[e^{is\sum_{i=1}^{N} Z_i}\right] = \exp\left[\int_{-\infty}^{\infty}(e^{isy}-1)\lambda F_Z(dy)\right].$$

Proof. Let Φ_Z be the common characteristic function of Z_n. By conditioning on the number of jumps and then using independence we get for any $s \in \mathbb{R}$,

$$\begin{aligned}
\Phi_X(s) &= \mathbb{E}\left[e^{is\sum_{i=1}^{N} Z_i}\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{is(Z_1+\cdots+Z_N)} \mid N=n\right] \mathbf{P}(N=n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{is(Z_1+Z_2+\cdots+Z_n)}\right] e^{-\lambda}\frac{\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{[\lambda\Phi_Z(s)]^n}{n!}.
\end{aligned}$$

That is

$$\Phi_X(s) = \exp\left[\lambda(\Phi_Z(s)-1)\right]. \tag{1.1}$$

Now with $\Phi_Z(s) = \int_{-\infty}^{\infty} e^{isy} F_Z(dy)$ it follows that

$$\Phi_X(s) = \exp\left[\lambda\left(\int_{-\infty}^{\infty} e^{isy}F_Z(dy) - 1\right)\right].$$

Using $\int_{-\infty}^{\infty} F_Z(dy) = 1$ we get that

$$\Phi_X(s) = \exp\left[\lambda\int_{-\infty}^{\infty}\left(e^{isy}-1\right)F_Z(dy)\right],$$

so the proof is complete. □

Now from (1.1), above, it can be easily seen that

$$\Phi_X(s) = \left[\exp\left[\frac{\lambda}{n}\left(\Phi_Z(s)-1\right)\right]\right]^n, \qquad s \in \mathbb{R},$$

implying that the compound Poisson distribution is infinitely divisible with each division following a $CP(\frac{\lambda}{n}, F_Z)$.

Remark 1.3 *We invariably write $\Phi_X(s)$ and $\Phi_F(s)$ to mean the characteristic function of the random variable X and characteristic function of its distribution F. This rises no confusion since there is a one–one correspondence between characteristic functions and distributions.*

Example 1.5 Inverse Gaussian (IG) random variables

If F has underlying random variable $X \sim IG(\mu, \theta)$ then for its density given by

$$f(x) = \frac{\mu}{\sqrt{2\pi} x^{3/2}} \exp\left\{\mu\theta - \frac{1}{2}(\frac{\mu^2}{x} + \theta^2 x)\right\}, \qquad x > 0 \text{ and } \mu,\ \theta > 0,$$

the characteristic function can be obtained as:

$$\begin{aligned} \Phi_F(s) &= \int_0^\infty e^{isx} \frac{\mu}{\sqrt{2\pi} x^{3/2}} \exp\left\{\mu\theta - \frac{1}{2}(\frac{\mu^2}{x} + \theta^2 x)\right\} dx \\ &= \exp\left\{ -\mu(\sqrt{-2is + \theta^2} - \theta)\right\}, \qquad s \in \mathbb{R}. \end{aligned} \tag{1.2}$$

Hence we can write

$$\Phi_F(s) = \left\{e^{\left[-\frac{\mu}{n}(\sqrt{-2is+\theta^2}-\theta)\right]}\right\}^n, \qquad s \in \mathbb{R},$$

and therefore recognize $F^{1/n}$ as the distribution with underlying random variable $Y \sim IG(\frac{\mu}{n}, \theta)$ having the characteristic function

$$\Phi_{F^{1/n}}(s) = e^{\left[-\frac{\mu}{n}(\sqrt{-2is+\theta^2}-\theta)\right]}, \qquad s \in \mathbb{R}.$$

Then $\Phi_F(s) = \left[\Phi_{F^{1/n}}(s)\right]^n$ and hence $F = (F^{1/n})^{*n}$, which implies by Proposition 1.2 that inverse Gaussian random variables are infinitely divisible.

1.4 Important Results Concerning IDD's

Now we intend to discuss various results concerning IDD's which are related with our work. The first one tells us what happens when we add two IDD's (or consider the convolution of two IDD measures; see Appendix A).

Theorem 1.2 *The sum of two infinitely divisible independent random variables is itself infinitely divisible.*

The proof results from a similar argument used for Lévy processes. The result implies that a finite sum of IDD's is itself IDD.

With this proposition and the examples of IDD's discussed above we now construct a new IDD, which is in fact the corner stone of the whole thesis.

Let $X = X_1 + X_2$, where $X_1 \sim N(\eta, \sigma^2)$ and $X_2 \sim CP(\lambda, F_Z)$ are independent. Then

$$\Phi_X(s) = \exp\left[i\eta s - \frac{1}{2}s^2\sigma^2 + \int_{-\infty}^{\infty} \lambda(e^{isy} - 1)F_Z(dy)\right], \qquad s \in \mathbb{R}. \tag{1.3}$$

By the above definition, Example 1.1, Proposition 1.3 and Theorem 1.1, X is infinitely divisible. Hence IDD's can be constructed by convolution of Gaussian and compound Poisson random variables.

The expression in (1.3) is close to the expression in the celebrated *Lévy–Khinchine formula.* This is further explored in the following section.

1.4.1 The Lévy measure

Before formally stating the *Lévy–Khinchine formula* it is necessary to introduce a special measure of important significance. This central element of the theory is known as *the Lévy measure.*

Let ν be a Borel measure defined on $\mathbb{R} \setminus \{0\}$. We say that ν is a Lévy measure if

$$\int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty\,, \tag{1.4}$$

or equivalently

$$\int_{|x|\leq 1} |x|^2\nu(dx) < \infty \qquad \text{and} \qquad \int_{|x|\geq 1} \nu(dx) < \infty. \tag{1.5}$$

Since $(|x|^2 \wedge \epsilon) \leq (|x|^2 \wedge 1)$ for all $0 < \epsilon \leq 1$ it follows that

$$\int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge \epsilon\right) \nu(dx) \leq \int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge 1\right) \nu(dx)\,, \qquad \text{if} \quad 0 < \epsilon \leq 1\,,$$

and hence from (1.4) it follows that

$$\int_{\mathbb{R}\setminus\{0\}} \left(|x|^2 \wedge \epsilon\right) \nu(dx) < \infty \qquad \Rightarrow \qquad \nu\left[(-\epsilon, \epsilon)^c\right] < \infty \qquad \text{for} \qquad 0 < \epsilon \leq 1\,. \tag{1.6}$$

So we have

$$\int_{|x|\leq\epsilon} |x|^2\nu(dx) < \infty \quad \text{and} \quad \int_{|x|\geq\epsilon} \nu(dx) < \infty, \qquad \text{if} \quad 0 < \epsilon \leq 1. \tag{1.7}$$

Remark 1.4 [About the Lévy measure]

[1] With the convention $\nu(\{0\}) = 0$ one can extend ν on $\mathbb{R}$.

[2] Any finite measure on $\mathbb{R} \setminus \{0\}$ is a Lévy measure but a Lévy measure on $\mathbb{R} \setminus \{0\}$ need not be finite. ν being a Radon measure, it is $\sigma-$ finite.

[3] Alternative characterizations of Lévy measures can be found in the literature. One of the most popular is

$$\int_{\mathbb{R}\setminus\{0\}} \left(\frac{|x|^2}{1+|x|^2}\right) \nu(dx) < \infty \tag{1.8}$$

It can be shown that (1.4) and (1.8) are equivalent. This latter form is useful in different computations, e.g. in estimation.

We will gain more insight on Lévy measures when we study the results for Lévy process using those for IDD's.

With the characteristic function in (1.3), for the convolution of two independent Gaussian and compound Poisson random variables and its relation to the notion of the Lévy measure, we are now in a position to state the central formula of the theory of Lévy processes.

1.4.2 Lévy–Khinchine formula

Theorem 1.3 *$F \in M_1(\mathbb{R})$ is infinitely divisible if there exists scalars $a, b \in \mathbb{R}$ and a Lévy measure ν on $\mathbb{R} \setminus \{0\}$ such that for all $s \in \mathbb{R}$:*

$$\Phi_F(s) = \exp\left[ias - \frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx)\right], \tag{1.9}$$

where $\mathbb{I}$ is an indicator function (see Definition A.2). Conversely any mapping of the above form is the characteristic function of an infinitely divisible probability measure on $\mathbb{R}$

A detailed proof can be found, for example, in, [16], [1], or [5]. We prefer the proof in [1] because of its constructive nature.

Remark 1.5 [On the Lévy–Khinchine formula]

[1] It represents all infinitely divisible random variables arising through the interplay between Gaussian and Poisson distributions.

[2] The "cut-off" function inside the integral $C(x) = \mathbb{I}_{\{-1,1\}}(x)$ takes different forms in the literature. The only constraint one needs to take into account in choosing C is that it needs to ensure the integrability of the function

$$g_c(x) = e^{isx} - 1 - isx \cdot C(x), \qquad x \in \mathbb{R},$$

with respect to ν, for each $s \in \mathbb{R}$.

[3] While choosing a cut-off function C it is sufficient to assume that $C(x) = 1 + O(|x|)$. With this restriction we see that the integral in (1.9) will never yield any term involving s^2 which means that the parameter b^2 is invariant of the choice of C. Also with the same restriction on C, the integral in (1.9) always yield a term involving is. This means that the parameter a should be adjusted with the choice of C. And as we saw before the definition of the Lévy measure is independent of the choice of C, so ν is also invariant of the choice of C. Thus relative to the choice of C the members of the triplet (a_c, b^2, ν) are called the characteristics of the infinitely divisible distribution F (or infinitely divisible random variable X)

[4] To use a cut-off function $C(x)$ other than $\mathbb{I}_{\{-1,1\}}(x)$ we need the following adjustments in the Lévy–Khinchine formula:

$$\Phi_F(s) = \exp\left[ia_c s - \frac{1}{2}s^2 b^2 + \int_{\mathbb{R}\setminus\{0\}} \left[e^{isx} - 1 - isx \cdot C(x)\right] \nu(dx)\right] \tag{1.10}$$

To find the relation of a_c to a we simply divide (1.10) by (1.9):

$$1 = \exp\left[is(a_c - a) + \int_{\mathbb{R}\setminus\{0\}} -isx\left[C(x) - \mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx)\right],$$

which implies that

$$0 = is(a_c - a) - is\int_{\mathbb{R}\setminus\{0\}} x\left[C(x) - \mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx),$$

and in turn that

$$a_c - a = \int_{\mathbb{R}\setminus\{0\}} x\left[C(x) - \mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx).$$

That is

$$a_c = a + \int_{\mathbb{R}\setminus\{0\}} x\left[C(x) - \mathbb{I}_{\{-1,1\}}(x)\right] \nu(dx). \tag{1.11}$$

[5] If ν satisfies the condition $\int_{|x|\leq 1} |x|\nu(dx) < \infty$, in addition to the usual conditions, we see that the integral in (1.11) exists with $C(x) \equiv 0$. Then the corresponding a_0 is known as the drift of F and is given by:

$$a_0 = a - \int_{\mathbb{R}\setminus\{0\}} x\mathbb{I}_{\{-1,1\}}(x)\nu(dx).$$

With this choice of $c \equiv 0$ and the corresponding a_c expressed as a_0, (1.10) turns into:

$$\begin{aligned}\Phi_F(s) &= \exp\Big[is\{a - \int_{\mathbb{R}\setminus\{0\}} x\,\mathbb{I}_{\{-1,1\}}(x)\,\nu(dx)\} \\ &\quad -\frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{isx}-1\right)\nu(dx)\Big] \\ &= \exp\Big[is\{a - \int_{\{\{|x|<1\}\setminus\{0\}\}} x\,\nu(dx)\} \\ &\quad -\frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{isx}-1\right)\nu(dx)\Big]. \qquad (1.12)\end{aligned}$$

This is the Lévy–Khinchine representation of F in terms of drift a_0, that is with respect to the characteristic triplet (a_0, b^2, ν).

Also if ν satisfies $\int_{|x|\geq 1} |x|\nu(dx) < \infty$ then we see that the integral in (1.11) exists with $C(x) \equiv 1$. Then the corresponding a_1 is known as the centre of F which is given by:

$$a_1 = a + \int_{\mathbb{R}\setminus\{0\}} x\left[1 - \mathbb{I}_{\{-1,1\}}(x)\right]\nu(dx).$$

With this choice of $C \equiv 1$ and the corresponding a_c expressed as a_1 (1.10) turns into:

$$\begin{aligned}\Phi_F(s) &= \exp\Big[is\Big\{a + \int_{\mathbb{R}\setminus\{0\}} x\left[1 - \mathbb{I}_{\{-1,1\}}(x)\right]\nu(dx)\Big\}\Big] \\ &\quad \exp\Big[-\frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \Big(e^{isx}-1-isx\Big)\nu(dx)\Big] \\ &= \exp\Big[is\Big\{a + \int_{|x|\geq 1} x\nu(dx)\Big\} - \frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \Big(e^{isx}-1-isx\Big)\nu(dx)\Big].\end{aligned} \qquad (1.13)$$

This is the Lévy–Khinchine representation of F in terms of center a_1, that is with respect to the characteristic triplet (a_1, b^2, ν).

It can be shown that the finiteness of $\int_{|x|>1} |x|\nu(dx)$ is equivalent to the finiteness of $\int_{\mathbb{R}} |x|dF(x)$ and that $a_1 = \int_{\mathbb{R}} xdF(x)$, the mean of F. Thus the centre and the mean are identical.

Equations (1.9), (1.12) and (1.13) explicitly shows the required adjustment one needs to carry out for choosing cut-off functions other than $\mathbb{I}_{\{-1,1\}}(x)$.

[6] As indicated in the discussion on the proof of the Lévy–Khinchine formula (for details see the proof in [1]) it is worth noting that all infinitely divisible distributions can be constructed as weak limits of the convolution between Gaussian with independent compound Poisson variables. This is precisely the reason why the expression in (1.3) is close to the expression of this Lévy–Khinchine formula.

The next theorem shows that in fact the compound Poisson distribution is sufficient for a weak approximation.

Theorem 1.4 *Any infinitely divisible probability measure can be constructed as the weak limit of a sequence of compound Poisson distributions.*

Proof. Let Φ be the characteristic function of an arbitrary infinitely divisible probability measure F, so that $\Phi^{1/n}$ is the characteristic function of $F^{1/n}$. Then for each $n \in \mathbb{N}$ we may define

$$\Phi_n(s) = \exp\left[n\{\Phi^{1/n}(s) - 1\}\right] = exp\left[\int_{\mathbb{R}} (e^{isx} - 1)nF^{1/n}(dx)\right], \qquad s \in \mathbb{R}.$$

So that Φ_n is the characteristic function of a compound Poisson distribution (see Proposition 1.3). We then have

$$\begin{aligned}\Phi_n(s) &= \exp\left[n\left(e^{(1/n)log[\Phi(s)]} - 1\right)\right] \\ &= \exp\left[\log\Phi(s) + n\cdot O(\frac{1}{n})\right] \longrightarrow \Phi(s) \quad \text{as} \quad n \to \infty,\end{aligned}$$

where log is the principal value of the logarithm. Then the result follows by Glivenko's theorem (see Appendix A). □

In a more general framework we have the following result.

Proposition 1.4 *If $\{F_n\}$ is a sequence of infinitely divisible distributions and $F_n \longrightarrow F$, then F is infinitely divisible, i.e. weak limits of sequences of infinitely divisible probability measures are infinitely divisible.*

For a proof we refer to [1].

We saw that the Lévy–Khinchine formula is related to the characteristic function of an infinitely divisible distribution F (or an F distributed random variable). It can be expressed, using two parameters and a measure, as an exponential function with a complex exponent, i.e.

$$\Phi_F(s) = e^{\Psi(s)} \quad \text{where} \quad \Psi : \mathbb{R} \longrightarrow \mathbb{C}.$$

The complex function Ψ is known as the characteristic exponent or Lévy exponent of F (or an F distributed random variable). Some authors prefer referring to it as the *Lévy exponent.*

Since we know that $|\Phi_F(s)| \leq 1$ (see Appendix A), then using $\Psi = Re(\Psi) + iIm(\Psi)$ we have

$$\begin{aligned} \left|\Phi_F(s)\right| &= \left|e^{\Psi(s)}\right| = \left|e^{Re(\Psi)+iIm(\Psi)}\right| \\ &= \left|e^{Re(\Psi)}\right|\left|e^{iIm(\Psi)}\right| = e^{Re(\psi)}1 \leq 1. \end{aligned}$$

Hence $e^{Re(\Psi)} \leq 1$ implies that $Re(\Psi) \leq 0$, that is the characteristic exponent should always have a non positive real part.

Chapter 2

Lévy Processes

With the tools and ideas discussed for IDD's we are now in a position to relate their properties with Lévy processes. Then we discuss some constructions involving Lévy processes.

2.1 Basic Ideas

The first result is at the core of the relation of Lévy processes and IDD's.

Theorem 2.1 *If* $X = \{X_t\,;\, t \geq 0\}$ *is a Lévy process, then* X_t *is infinitely divisible for each* $t \geq 0$.

Proof. For each $n \in \mathbb{N}$, we can write

$$X_t = Y_1^n(t) + Y_2^n(t) + \cdots + Y_n^n(t),$$

where each

$$\begin{aligned} Y_k^n(t) &= X_{\frac{kt}{n}} - X_{\frac{(k-1)t}{n}}, & k = 1, 2, \cdots, n, \\ &\overset{D}{=} X_{(\frac{kt}{n} - \frac{kt}{n} + \frac{t}{n})}, & \text{by [L2]-(ii) of Definition1.1}, \\ &= X_{\frac{t}{n}}. \end{aligned}$$

The last term in the above equality is independent of k, which shows that for all k, the $Y_k^n(t)$'s are iid with the common distribution $X_{\frac{t}{n}}$. Hence

$$\Phi_{X_t}(s) = \left[\Phi_{X_{\frac{t}{n}}}(s)\right]^n, \qquad s \in \mathbb{R},$$

which shows that X_t, for each $t \geq 0$, is infinitely divisible. □

Remark 2.1

[1] The term infinitely decomposable would be more appropriate than infinitely divisible.

[2] Here n can take any value in $\mathbb{N}$, which has an infinite cardinality. Hence the name infinitely divisible.

By the above theorem we can write that $\Phi_{X_t}(s) = e^{\Psi(t,s)}$, for each $t \geq 0$ and $s \in \mathbb{R}$, where $\Psi(t, \cdot)$ is a Lévy exponent.

We now show that $\Psi(t, s) = t\Psi(s)$, for each $t \geq 0$ and $s \in \mathbb{R}$. To prove this result we need the following lemma.

Lemma 2.1 *If $X = \{X_t\,;\, t \geq 0\}$ is stochastically continuous then the map $t \longrightarrow \Phi_{X_t}(s)$ is continuous for each $s \in \mathbb{R}$.*

Theorem 2.2 *If X is a Lévy process then $\mathbb{E}[e^{isX_t}] = \Phi_{X_t}(s) = e^{t\Psi(s)}$, for each $t \geq 0$ and $s \in \mathbb{R}$, where Ψ is the Lévy exponent of X_1.*

Proof. Suppose that X is a Lévy process and that $\Phi_{X_t}(s)$, for each $s \in \mathbb{R}$ and $t \geq 0$ is the characteristic function of X_t. Then from Definition 1.1-L2, we have for all $h \geq 0$,

$$\begin{aligned}
\Phi_{X_{t+h}}(s) &= \mathbb{E}\left(e^{isX_{t+h}}\right) \\
&= \mathbb{E}\left[e^{is(X_{t+h}-X_t)}\right]\mathbb{E}\left(e^{isX_t}\right), && \text{using [L2]-(i)} \\
&= \mathbb{E}\left[e^{is(X_h)}\right]\mathbb{E}\left(e^{isX_t}\right) && \text{using [L2]-(ii)} \\
&= \Phi_{X_h}(s)\cdot\Phi_{X_t}(s). && (2.1)
\end{aligned}$$

Now by L1 we have that

$$\Phi_{X_0}(s) = 1. \tag{2.2}$$

Also, by L3, X_t is stochastically continuous and so we can apply Lemma 2.1 to see that the map $t \longrightarrow \Phi_{X_t}(s)$ is continuous. Together with the multiplicative property in (2.1) this implies that $\Phi_{X_t}(s)$ is an exponential function in t. Hence the unique continuous solution of (2.1) and (2.2) is given by

$$\Phi_{X_t}(s) = e^{tK(s)}, \qquad s \in \mathbb{R},$$

where $K : \mathbb{R} \to \mathbb{C}$. Now by Theorem 2.1, X_1 is infinitely divisible. Hence

$$\Phi_{X_1}(s) = e^{K(s)}, \qquad s \in \mathbb{R}, \tag{2.3}$$

which implies that K is the Lévy exponent of X_1, completing the proof. □

Remark 2.2

[1] Recalling the definition of cumulant generating function of a random variable from Appendix A, we see that Ψ in the above theorem is the cumulant generating function of X_1 i.e. $\Psi = K_{X_1}$ and that it varies linearly in t, i.e.

$$K_{X_t} = tK_{X_1} = t\Psi.$$

The distribution of X_t is therefore determined by the knowledge of the distribution of X_1. Thus the only degree of freedom we have in specifying a Lévy process is to specify the distribution of X_t for a single period say $t = 1$.

[2] The Lévy exponent of X_1 is simply the cumulant of X_1 and Theorem 2.2 is useful as the definition of the characteristic function of the Lévy process $X = \{X_t\,;\, t \geq 0\}$.

We now have the complete expression of the Lévy–Khinchine formula for the Lévy process $X = \{X_t\,;\, t \geq 0\}$ as:

$$\mathbb{E}[e^{isX_t}] = \exp\left\{t\left[ias - \frac{1}{2}s^2b^2 + \int_{\mathbb{R}\setminus\{0\}} \left[e^{isx} - 1 - isx\mathbb{I}_{\{-1,1\}}(x)\right]\nu(dx)\right]\right\}. \tag{2.4}$$

Comparing (1.9) with (2.4) we observe that the latter is simply a version of the former corresponding to $t = 1$, which leads to the following more convincing definition of the Lévy measure of a Lévy process.

The characteristic triplet of the Lévy process is just the characteristic triplet of the infinitely divisible random variable X_1. It can be shown that this correspondence is unique. Thus given a Lévy process there corresponds a unique IDD which is the distribution of X_1. It follows that corresponding to every infinitely divisible distribution there exists a Lévy process so that the characteristics of the process evaluated at $t = 1$ coincide with the characteristics of the IDD.

Definition 2.1 *[Lévy measure of a Lévy process] For a Lévy process $X = \{X_t\,;\, t \geq 0\}$ on $\mathbb{R}$, the measure ν on $\mathbb{R}$ defined by:*

$$\nu(A) = \mathbb{E}\left[\sharp\{t \in [0,1] \mid \Delta X_t \neq 0, \Delta X_t \in A\}\right]\,, \qquad A \in \mathcal{B}(\mathbb{R}), \tag{2.5}$$

is called the Lévy measure of X. Here $\nu(A)$ is the expected number, per unit time, of the jumps with size in set A (see [3]).

We now see what happens when two independent Lévy processes are added.

Theorem 2.3 *The sum of two independent Lévy processes is also a Lévy process.*

We need the following lemma to prove the above theorem. This lemma is known in the literature as Kac's theorem (see [1]).

Lemma 2.2 *The random variables $X_1, \cdots, X_n$ are independent if and only if*

$$\mathbb{E}\left(e^{i\sum_{j=1}^{n} u_j X_j}\right) = \Phi_{X_1}(u_1)\cdots\Phi_{X_n}(u_n), \qquad \textit{for all} \;\; u_1, \cdots, u_n \in \mathbb{R}. \tag{2.6}$$

Proof of Theorem 2.3. Let X_t and Y_t be two independent Lévy processes and consider

$$Z_t = X_t + Y_t\,, \qquad t \geq 0\,.$$

By Definition 1.1-[L1], we have that $Z_0 = X_0 + Y_0 = 0$, a.s.. To prove the independence of the increments consider for any $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_n < \infty$:

$$\begin{aligned}
&\mathbb{E}\Big[e^{i\left\{\sum_{k=0}^{n-1} u_{k+1}(Z_{t_{k+1}} - Z_{t_k})\right\}}\Big] \\
&= \mathbb{E}\Big[e^{i\left\{\sum_{k=0}^{n-1} u_{k+1}(X_{t_{k+1}} - X_{t_k})\right\}}\Big]\,\mathbb{E}\Big[e^{i\left\{\sum_{k=0}^{n} u_{k+1}(Y_{t_{k+1}} - Y_{t_k})\right\}}\Big]\,, && \text{by independence}\,,\\
&= \prod_{k=0}^{n-1} \Phi_{X_{t_{k+1}} - X_{t_k}}(u_{k+1}) \prod_{k=0}^{n-1} \Phi_{Y_{t_{k+1}} - Y_{t_k}}(u_{k+1})\,, && \text{by [L2]-(i) and (2.6)}\,,\\
&= \prod_{k=0}^{n-1} \Phi_{X_{t_{k+1}} - X_{t_k}}(u_{k+1})\,\Phi_{Y_{t_{k+1}} - Y_{t_k}}(u_{k+1})\,, && \text{for all } u_1, \cdots, u_n \in \mathbb{R}\,,\\
&= \prod_{k=0}^{n-1} \Phi_{X_{t_{k+1}} - X_{t_k} + Y_{t_{k+1}} - Y_{t_k}}(u_{k+1})\,, && \text{by the independence of } \{X_t\} \text{ and } \{Y_t\}\,,\\
&= \prod_{k=0}^{n-1} \Phi_{Z_{t_{k+1}} - Z_{t_k}}(u_{k+1})\,.
\end{aligned}$$

Considering (2.6) again this implies that Z_t has independent increments.

Now to show that Z_t has stationary increments we proceed as follows. Let $\Delta X_j = X_{t_{j+1}} - X_{t_j}$ and $\Delta Y_j = Y_{t_{j+1}} - Y_{t_j}$. Since X_t and Y_t have stationary increments it follows that:

$$F_{\Delta X_j} = F_{X_{(t_{j+1}-t_j)}} \quad \text{and} \quad F_{\Delta Y_j} = F_{Y_{(t_{j+1}-t_j)}}.$$

Now we observe that $Z_{t_{j+1}} - Z_{t_j} \stackrel{def}{=} \Delta Z_j = \Delta X_j + \Delta Y_j$ and that ΔX_j and ΔY_j are independent. Hence by the definition of convolution (see Appendix A).

$$\begin{aligned} F_{(Z_{t_{j+1}}-Z_{t_j})} &= F_{\Delta Z_j} = F_{(\Delta X_j + \Delta Y_j)} = F_{\Delta X_j} * F_{\Delta Y_j} \\ &= F_{X_{(t_{j+1}-t_j)}} * F_{Y_{(t_{j+1}-t_j)}} = F_{(X_{(t_{j+1}-t_j)} + Y_{(t_{j+1}-t_j)})} = F_{Z_{(t_{j+1}-t_j)}}, \end{aligned}$$

which shows that Z_t has stationary increments.

Finally the stochastic continuity follows from the elementary inequality of probability:

$$\mathbf{P}(| Z_t - Z_s |> a) \le \mathbf{P}(| X_t - X_s |> \frac{a}{2}) + \mathbf{P}(| Y_t - Y_s |> \frac{a}{2}) \longrightarrow 0 \quad \text{as} \quad t \to s,$$

since X_t and Y_t are Lévy processes. Hence Z_t is a Lévy process. □

Now because of the one-to-one correspondence between Lévy processes and IDD's, as discussed above, the following result is immediate.

Corollary 2.1 *The sum of two infinitely divisible random variables is itself infinitely divisible.*

2.2 Examples of Lévy Processes

This section presents some important examples of Lévy processes and their properties.

2.2.1 The Poisson Process

From Appendix A we recall that if N is a Poisson random variable with parameter λ then its characteristic function is given by:

$$\Phi_N(s) = e^{\lambda[e^{is}-1]}, \qquad s \in \mathbb{R}. \tag{2.7}$$

We now define the Poisson process.

Definition 2.2 *Let $\{\tau_i\}_{i\geq 1}$ be a sequence of independent exponential random variables with parameter λ and let $T_n = \sum_{i=1}^{n} \tau_i$. The processes N_t defined by*

$$N_t = \sum_{n\geq 1} 1_{[t\geq T_n]} \tag{2.8}$$

is called a Poisson process with intensity λ.

The Poisson process therefore counts the number of random times T_n which occur between 0 and t, where $\{T_n - T_{n-1}\}_{n\geq 1}$ is an i.i.d. sequence of exponential variables. Why this process is called Poisson is better understood from the following proposition.

Proposition 2.1 *If $\{\tau_i\}_{i\geq 1}$ are independent and identically distributed exponential random variables with parameter λ then for any $t > 0$ the random variable*

$$N_t = \inf\{n \geq 1\,;\, \sum_{i=1}^{n} \tau_i > t\} - 1 \tag{2.9}$$

follows a Poisson distribution with parameter λt:

$$\mathbf{P}\{N_t = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \qquad \forall n \in \mathbb{N}. \tag{2.10}$$

A nice intuitive proof can be found e.g. in [5].

From (2.7) and (2.10) it follows that the characteristic function of a Poisson process is given by

$$\mathbb{E}[e^{isN_t}] = \exp\{\lambda t(e^{is} - 1)\}, \qquad \forall s \in \mathbb{R}. \tag{2.11}$$

The following proposition lists many important properties of Poisson processes.

Proposition 2.2 *Let N_t be a Poisson process. Then*

[1] for any $t > 0$, N_t is almost surely finite.

[2] for any ω, the sample path $t \rightarrow N_t(\omega)$ is piecewise constant and increases by jumps of size 1.

[3] the sample paths $t \rightarrow N_t$ are cadlag.

[4] for any $t > 0$, $\quad N_{t^-} = N_t$ a.s.

[5] N is continuous in probability, i.e. $\forall t > 0, N_s \xrightarrow{\mathbf{P}} N_t$ *when* $s \to t$.

[6] N has independent increments, i.e. for any $t_1 < \cdots < t_n$,

ΔN_{t_j}*'s are independent random variables.*

[7] the increments of N are homogeneous (stationary), i.e.

for any $t, h > 0\,, N_{t+h} - N_t \stackrel{D}{=} N_h$.

[8] N has the Markov property i.e.

$\forall t, h > 0\,,\ \mathbb{E}[f(N_{t+h}) \mid N_s\,, s \leq t] = \mathbb{E}[f(N_{t+h}) \mid N_t]$.

Proofs can be found in any standard book on this subject, e.g. [5].

Proposition 2.3 *The Poisson process is a Lévy process.*

One possible proof follows from Example 1.3 and the discussion after (2.4). Here we recall (2.10) to note that the characteristic of this Lévy process is just the characteristic of the infinitely divisible random variable N_1. Another proof follows as a corollary of the similar result presented for the compound Poisson process in the next section.

2.2.2 Compound Poisson process

The definition of a compound Poisson process that follows uses Definition 1.3 of a compound Poisson random variable.

Definition 2.3 *A compound Poisson process with intensity* $\lambda > 0$ *is a stochastic process* $X = \{X_t\}$ *defined as*

$$X_t = \begin{cases} \sum_{n=1}^{N_t} Z_n = Z_1 + \cdots + Z_{N_t} & \text{if } N_t > 0, \\ 0 & \text{if } N_t = 0, \end{cases} \tag{2.12}$$

where jump sizes Z_n *are iid with common distribution* F_Z *and* $N = \{N_t, t > 0\}$ *is a Poisson process with intensity* λ*, independent of* $\{Z_n\}_{n \geq 1}$.

Remark 2.3

[1] Again the sample paths of X are cadlag, piecewise constant functions on finite intervals with jump-discontinuities at random times T_n. However this time the size of the jumps is itself random and the jump at T_n can take any value in the range of the random variable Z_n.

[2] The jump times $\{T_n\}_{n\geq 1}$ of X are the jump times of the Poisson process N, they can be expressed as the partial sums of the independent exponential random variables with mean $\frac{1}{\lambda}$.

[3] The Poisson process itself can be seen as a compound Poisson process on $\mathbb{R}$ such that $Z_n \equiv 1$ for all n, which explains the origin of the term compound Poisson in the definition.

A simple derivation gives the characteristic function of a compound Poisson process:

$$\mathbb{E}[e^{isX_t}] = \exp\{t\lambda \int_{-\infty}^{\infty} (e^{isx} - 1)F_Z(dx)\}, \qquad \forall s \in \mathbb{R}, \tag{2.13}$$

where λ denotes the jump intensity and F the jump size distribution.

Theorem 2.4 *The compound Poisson process X in Definition 2.3 is a Lévy process, where N is a Poisson process with intensity λ and the Z_i's are i.i.d. random variables independent of N.*

Intuitively the proof follows from Example 1.4 and the relation between IDD's and Lévy processes. However a detailed proof follows.

Proof. $N_0 = 0$ a.s., by the definition of a Poisson process, which implies that $X_0 = 0$ a.s. For $0 = t_0 < t_1 < t_2 < \cdots < t_n$

$$\begin{aligned}
&\mathbb{E}\Big[e^{i\sum_{j=0}^{n-1}\left(u_{j+1}\sum_{k=N_{t_j}+1}^{N_{t_{j+1}}} Z_k\right)}\Big] \\
&= \sum_{i_1=0}^{\infty}\cdots\sum_{i_n=0}^{\infty}\mathbb{E}\Big[e^{i\sum_{j=0}^{n-1}\left(u_{j+1}\sum_{k=N_{t_j}+1}^{N_{t_{j+1}}} Z_k\right)} \mid N_{t_1} = i_1, N_{t_2} = i_1 + i_2, \cdots, \\
&\qquad N_{t_n} = i_1 + \cdots + i_n\Big]\mathbf{P}\{N_{t_1} = i_1, N_{t_2} = i_1 + i_2, \cdots, N_{t_n} = i_1 + \cdots + i_n\} \\
&= \sum_{i_1=0}^{\infty}\cdots\sum_{i_n=0}^{\infty}\Phi_Z^{i_1}(u_1)\cdots\Phi_Z^{i_n}(u_n)\mathbf{P}\{N_{t_1} - N_{t_0} = i_1, \cdots, N_{t_n} - N_{t_{n-1}} = i_n\} \\
&= \sum_{i_1=0}^{\infty}\cdots\sum_{i_n=0}^{\infty}\Phi_Z^{i_1}(u_1)\cdots\Phi_Z^{i_n}(u_n)\mathbf{P}\{N_{t_1} - N_{t_0} = i_1\}\cdots\mathbf{P}\{N_{t_n} - N_{t_{n-1}} = i_n\}.
\end{aligned}$$

(2.14)

Now, by independence we also have

$$\begin{aligned}
&\mathbb{E}\Big[e^{i\left(u_1 \sum_{k=N_{t_0}+1}^{N_{t_1}} Z_k\right)}\Big] \cdots \mathbb{E}\Big[e^{i\left(u_n \sum_{k=N_{t_{n-1}}+1}^{N_{t_n}} Z_k\right)}\Big] \\
&= \sum_{i_1=0}^{\infty} \cdots \sum_{i_n=0}^{\infty} \mathbb{E}\Big[e^{i\left(u_1 \sum_{k=N_{t_0}+1}^{N_{t_1}} Z_k\right)} | N_{t_1} - N_{t_0} = i_1\Big] \mathbf{P}\{N_{t_1} - N_{t_0} = i_1\} \cdots \\
&\qquad \mathbb{E}\Big[e^{i\left(u_n \sum_{k=N_{t_{n-1}}+1}^{N_{t_n}} Z_k\right)} | N_{t_n} - N_{t_{n-1}} = i_n\Big] \mathbf{P}\{N_{t_n} - N_{t_{n-1}} = i_n\} \\
&= \sum_{i_1=0}^{\infty} \cdots \sum_{i_n=0}^{\infty} \Phi_Z^{i_1}(u_1) \cdots \Phi_Z^{i_n}(u_n) \mathbf{P}\{N_{t_1} - N_{t_0} = i_1\} \cdots \mathbf{P}\{N_{t_n} - N_{t_{n-1}} = i_n\}. \qquad (2.15)
\end{aligned}$$

Hence from (2.14) and (2.15) we get that $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.

Now to show stationarity, let $0 \leq s < t$, then

$$\begin{aligned}
\mathbb{E}\Big[e^{iu(X_t - X_s)}\Big] &= \mathbb{E}\Big[e^{i\left(u \sum_{k=N_s+1}^{N_t} Z_k\right)}\Big] \\
&= \sum_{j=0}^{\infty} \mathbb{E}\Big[e^{i\left(u \sum_{k=N_s+1}^{N_t} Z_k\right)} | N_t - N_s = j\Big] \mathbf{P}\{N_t - N_s = j\} \\
&= \sum_{j=0}^{\infty} \Phi_Z^j(u) \mathbf{P}\{N_t - N_s = j\} \\
&= \sum_{j=0}^{\infty} \Phi_Z^j(u) \mathbf{P}\{N_{t-s} = j\}, \qquad \text{by Definition 2.2,} \\
&= \mathbb{E}\Big[e^{i\left(u \sum_{k=1}^{N_{t-s}} Z_k\right)}\Big] = \mathbb{E}\Big[e^{i(uX_{t-s})}\Big],
\end{aligned}$$

which shows that $X_t - X_s \stackrel{D}{=} X_{t-s}$.

Finally to show the stochastic continuity, consider for any $a > 0$ and $t > 0$

$$\begin{aligned}
\mathbf{P}\{| X_t |> a\} &= \sum_{i=1}^{\infty} \mathbf{P}\{| Z_1 + Z_2 + \cdots + Z_i |> a \big| N_t = i\} \mathbf{P}\{N_t = i\} \\
&= \sum_{i=1}^{\infty} \mathbf{P}\{| Z_1 + Z_2 + \cdots + Z_i |> a\} \mathbf{P}\{N_t = i\} \\
&\leq 1 - e^{-\lambda t} \longrightarrow 0, \quad \text{as} \quad t \longrightarrow 0,
\end{aligned}$$

where the last equality holds since Z_n and N are independent. Hence the compound Poisson process is a Lévy process. □

The converse statement of the above theorem is also true, provided we add a condition. This leads to the following more general result.

Theorem 2.5 $X = \{X_t\,;\, t \geq 0\}$ *is a compound Poisson process if and only if it is a Lévy process and its sample paths are piecewise constant functions.*

One implication of Theorem 2.5 was proved above. The proof for the converse implication, which is rather involved, can be found in [5].

Thus to conclude whether a general Lévy process is a compound Poisson process or not it is sufficient to observe whether its sample paths are piecewise constant. In fact the compound Poisson process is the only Lévy process with piecewise constant sample paths (considering the Poisson process as the simplest compound Poisson process). The following result is thus a simple corollary of Theorem 2.5.

Corollary 2.2 *The Poisson process is a Lévy process.*

Equation (2.13) is a special case of the Lévy–Khinchine formula in (1.9) or (2.4). Let us introduce a new measure ν defined by $\nu(A) = \lambda F(A)$. Then

$$F_Z(A) = \frac{\nu(A)}{\lambda}. \tag{2.16}$$

Here $A \in \mathcal{B}(\mathbb{R})$ (the collection of all Borel sets in $\mathbb{R}$) and $\nu(A)$ is the Lévy measure of the compound Poisson(λ, F_Z) process, then the Lévy–Khinchine representation becomes

$$\mathbb{E}[e^{isX_t}] = \exp\left\{t \int_{-\infty}^{\infty} (e^{isx} - 1)\nu(dx)\right\}, \qquad \forall s \in \mathbb{R}. \tag{2.17}$$

Since $\nu(\mathbb{R}) = \int_{-\infty}^{\infty} \lambda dF_Z(x) = \lambda \neq 1$, ν is a positive measure on $\mathbb{R}$ but not a probability measure.

2.2.3 Gaussian Lévy process

Definition 2.4 *A standard Brownian motion in* $\mathbb{R}$ *is a Lévy process* $B = \{B_t\,;\, t \geq 0\}$ *for which:*

[B1] $B_t \sim N(0, t)$ *for each* $t \geq 0$,

[B2] B *has continuous sample paths.*

In fact Brownian motion is the only Lévy process with continuous sample paths. From $[B1]$ and the definition of characteristic function (see Appendix A) it follows immediately that the characteristic function of Brownian motion is given by:

$$\Phi_{B_t}(s) = \exp\{-\frac{1}{2}ts^2\}, \qquad \text{for all } s \in \mathbb{R} \text{ and } t \geq 0. \tag{2.18}$$

Remark 2.4 The terminology "Brownian motion" may correspond to an arbitrary variance–covariance matrix where as the terminology "Wiener process", $\{W_t\,;\, t \geq 0\}$ is used only for standard Brownian motion with unit co-variance matrix that is $W_t = B_t \sim N(0,t)$.

Thus Brownian motion is by definition a Lévy process. Another more intuitive derivation of this fact follows from Example 1.1 and the one–one correspondence of IDD's with Lévy processes. The deterministic process $\{X_t = at\,;\, t \geq 0\}$ is known as the linear drift. The following proposition states that it is a deterministic Lévy process.

Proposition 2.4 *The simplest Lévy process is the linear drift.*

The proof follows directly from the Lévy–Khinchine representation of a Lévy process. Now by Theorem 2.3, $X_t = \{at + W_t\,;\, t \geq 0\}$ is also a Lévy process which is known as "Brownian motion with drift".

2.3 Random Measure of Lévy Processes With Threshold Exceeding Jumps

Since a Lévy process is cadlag, the number of jumps ΔX_s such that $|\Delta X_s| \geq \epsilon$, before some time t, has to be finite for all $\epsilon > 0$. Hence if $B \in \mathcal{B}(\mathbb{R})$, is bounded away from 0 (i.e. $0 \notin \bar{B}$, the closure of B), then for $t \geq 0$

$$N_t^B = \sharp\{s \in [0,t]\,;\, \Delta X_s \in B\} = J_X([0,t] \times B)\,, \tag{2.19}$$

is well defined and a.s. finite. The process N^B is clearly a counting process, called a counting process of B. It inherits the Lévy property from X. Since the Poisson process is the only non–trivial counting process which is Lévy then N_t^B is a Poisson process with a certain intensity $\nu^X(B) < \infty$. If B is a disjoint union of Borel sets B_i, then $N_t^B = \sum_i N_t^{B_i}$. Hence considering (2.5), $\nu^X(B) = \mathbb{E}(N_1^B) = \sum \mathbb{E}(N_1^{B_i}) = \sum \nu(B_i)$. ν is a Borel measure and as indicated in (1.6) it holds that $\nu(\mathbb{R} \setminus (-\epsilon, \epsilon)) < \infty$ for all $\epsilon > 0$. In particular ν is σ–finite (see Appendix A).

Now to discuss the term $J_X([0,t] \times B)$ in (2.19) we need the ideas of a random measure and Poisson random measure. However we keep our discussion about these two measures at an introductory level. For more details see [3] and [1].

From the definition of the Poisson process we recall that the jump times $T_1, T_2 \cdots$ form a random configuration of points on $[0, \infty)$ and that the Poisson process N_t counts the number of such points in the interval $[0,t]$.This counting procedure defines a measure on $[0, \infty)$.

Definition 2.5 *For any measurable set $A \subset \mathbb{R}^+$ a positive integer valued counting measure $M(\omega, \cdot)$ defined as*

$$M(\omega, A) = \sharp\{i \geq 1\,;\, T_i(\omega) \in A\}, \qquad \omega \in \Omega, \tag{2.20}$$

is a random measure.

The very first property of a Poisson process ensures that $M(A)$, for any bounded measurable set A, is almost surely finite. The intensity λ of the Poisson process determines the average value of the random measure M, i.e. $\mathbb{E}[M(A)] = \lambda|A|$ where $|A|$ is the Lebesgue measure of A. M is also known as a random jump measure associated to the Poisson process N. The Poisson process can be expressed in terms of the random measure M in the following way:

$$N_t(\omega) = M(\omega, [0,t]) = \int_{[0,t]} M(\omega, ds)\,.$$

The properties of the Poisson process described in Proposition 2.2, can be translated into properties of the measure M. Some of the important ones are as follows.

[1] For disjoint intervals $[t_1, t_1'], \cdots, [t_n, t_n']$, $M([t_k, t_k'])$ is the number of jumps of the Poisson process in $[t_k, t_k']$. It is a Poisson random variable with parameter $\lambda(t_k' - t_k)$. Generally for any measurable set A, $M(A)$ follows a Poisson distribution with parameter $\lambda|A|$, where $|A| = \int_A dx$ is the Lebesgue measure of A.

[2] For two disjoint intervals $[t_i, t_i']$ and $[t_j, t_j']$ where $i \neq j$, $M([t_i, t_i'])$ and $M([t_j, t_j'])$ are independent random variables.

A natural extension of this notion of random measure is the Poisson random measure, where $\mathbb{R}^+$ is replaced by any $E \subset \mathbb{R}$ and the Lebesgue measure by any Radon measure μ on E.

Definition 2.6 *Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, $E \subset \mathbb{R}$ and μ a given positive Radon measure on $(E, \mathcal{E})$. A Poisson random measure on E with intensity measure μ is an integer valued random measure*

$$M : \Omega \times E \to \mathbb{N}, \quad (\omega, A) \mapsto M(\omega, A)$$

such that:

[1] for almost all $\omega \in \Omega, M(\omega, \cdot)$ is an integer valued Radon measure on E, i.e. for any bounded measurable $A \subset E$, $M(A) < \infty$ is an integer valued random variable.

[2] for each measurable set $A \subset E, M(\cdot, A) = M(A)$ is a Poisson random variable with parameter $\mu(A)$, i.e.

$$\mathbf{P}\{M(A) = k\} = e^{-\mu(A)} \frac{[\mu(A)]^k}{k!}, \qquad \forall k \in \mathbb{N}. \tag{2.21}$$

[3] for disjoint measurable sets $A_1, \cdots, A_n \in \mathcal{E}$, the corresponding random variables $M(A_1), \cdots, M(A_n)$ are independent.

We state the following proposition, without proof, which ensures the existent of the Poisson random measure.

Proposition 2.5 *For any Radon measure μ on $E \subset \mathbb{R}$, there exists a Poisson random measure M on E with intensity μ.*

For a proof see [5].

It can be shown that to every cadlag process and in particular to every compound Poisson process $X = \{X_t\,;\, t \geq 0\}$ on $\mathbb{R}$ we can associate a random measure on $\mathbb{R} \times [0, \infty)$ describing the jumps of X. For every measurable set $B \subset \mathbb{R} \times [0, \infty)$

$$J_X(B) = J_X(A \times [t_1, t_1']) = \sharp\{t \in [t_1, t_1']\,;\, \Delta X_t \in A\}, \tag{2.22}$$

that is $J_X(A \times [t_1, t_1'])$:=number of jumps in X occurring between time t_1 to t_1' whose amplitude belongs to A.

The random measure J_X contains all information about the discontinuities (jumps) of X. It tells us when the jumps occur and how big they are. J_X does not give any information on the continuous part of X. It is easy to see, indeed we will see, that X has continuous sample paths if and only if $J_X = 0$ a.s. which implies that there are no jumps.
The following proposition shows that J_X is a Poisson random measure in the sense defined above.

Proposition 2.6 *Let $X = \{X_t\,;\, t \geq 0\}$ be a compound Poisson process with intensity λ and jump size distribution F. Its jump measure J_X is a Poisson random measure on $\mathbb{R} \times [0, \infty)$ with intensity measure $\mu(dx \times dt) = \nu(dx)dt = \lambda dF(x)dt$.*

For a proof see [5].

Thus to see the distribution of the number of jumps of size larger than a threshold ϵ, consider $A = [\epsilon, \infty)$ in (2.22). Then using the above proposition together with (2.21) we have

$$\begin{aligned}
\mathbf{P}\{J_X(A \times [t_1, t_1']) = k\} &= \exp\Big[-\int_A \int_{t_1}^{t_1'} \mu(dx \times dt)\Big] \frac{\Big[\int_A \int_{t_1}^{t_1'} \mu(dx \times dt)\Big]^k}{k!} \\
&= \exp\Big[-\int_A \int_{t_1}^{t_1'} \lambda F(dx)dt\Big] \frac{\Big[\int_A \int_{t_1}^{t_1'} \lambda F(dx)dt\Big]^k}{k!} \\
&= \exp\Big[-|t_1' - t_1| \int_A \nu(dx)\Big] \frac{\big[|t_1' - t_1| \int_A \nu(dx)\big]^k}{k!} \\
&= e^{[-|t_1'-t_1|\nu([\epsilon,\infty])]} \frac{\big[|t_1' - t_1|\nu([\epsilon, \infty])\big]^k}{k!}. \qquad (2.23)
\end{aligned}$$

Equation (2.23) helps us clarify one of the intuitive results involving Lévy process and its jump measure which we mentioned earlier. We now give its proof.

Proposition 2.7 *X has continuous sample paths if and only if $J_X = 0$ a.s. which implies that there are no jumps.*

Proof. $\mathbf{P}\{J_X(A \times [t_1, t_1']) = 0\} = 1$ implies, by (2.23), that

$$\exp\Big[-|t_1' - t_1|\nu([\epsilon, \infty])\Big] \frac{\big[|t_1' - t_1|\nu([\epsilon, \infty])\big]^0}{0!} = 1.$$

Which, for any choice of t_1', t_1 and ϵ, directly implies that $\nu(\epsilon,\infty)=0$, since $|t_1'-t_1|$ can not be zero. But then using equation (2.4) of the Lévy–Khinchine representation of a Lévy process we see that X is just a Brownian motion with drift. And so X has continuous sample paths.

On the other hand, if X has continuous sample paths then its Lévy measure, which controls all jumps, should be identically zero. That is $\nu(\epsilon,\infty)$ should be identically zero for any choice of ϵ. Then (2.23) implies for any positive integer k

$$\mathbf{P}\{J_X(A\times[t_1,t_1'])=k\}=0$$

and since 0^0 is numerically 1, for $k=0$ we get

$$\mathbf{P}\{J_X(A\times[t_1,t_1'])=0\}=1.$$

That is J_X is almost surely zero. Hence the proof is complete. □

2.4 Poisson Point Process

In this section we define the Poisson point process and relate it with the Poisson random measure given in Definition 2.6 and characterized by (2.22). This relation will help us avail an estimate of the Lévy measure of a subordinator through a construction. The statement of this construction appears in Theorem 3.1.

Definition 2.7 *The Poisson point process* $(\triangle_t)_{t\geq 0}$ *with intensity function* $g:(0,\infty)\to[0,\infty)$ *is a process such that*

$$N((t_1,t_1']\times(c,d])=\sharp\{t_1<t\leq t_1':\triangle_t\in(c,d]\}\sim Poi\left((t_1'-t_1)\int_c^d g(x)dx\right),\qquad(2.24)$$

$0\leq t_1<t_1'$, $(c,d]\subset(0,\infty)$. *It defines a Poisson counting measure on* $[0,\infty)\times(0,\infty)$.

Interpret $\triangle_s$ as a jump at time s, then it can be shown that $\sum_{0\leq s\leq t}\triangle_s$, the process performing all these jumps, is a Lévy process.

Note that according to this definition there is no restriction in choosing $c=\epsilon$ and $d=\infty$. Thus we have:

$$\mathbf{P}\{N((t_1,t_1']\times(\epsilon,\infty])=k\}=\exp\Big[-|t_1'-t_1|\int_\epsilon^\infty g(x)dx\Big]\frac{\big[|t_1'-t_1|\int_\epsilon^\infty g(x)dx\big]^k}{k!}.\qquad(2.25)$$

Equations (2.23) and (2.25) show that the distribution of the Poisson point process follows the distribution of the Poisson random measure, provided that the intensity function g of the Poisson point process is the density of the Lévy measure ν. This is where this process gets its name. In chapter three we will see that in deed g is the density of the Lévy measure. For more details on Poisson point processes and Poisson random measures we refer to [3].

Chapter 3

Special Lévy Processes and their Applications

In this chapter we discuss Lévy processes associated with special jump measures. Consider discarding those jumps which are less than a certain threshold. We see also how to get the density of such truncated jumps and how they are related to the compound Poisson (CP) process through a construction with the gamma process.

3.1 Distribution of Truncated Jumps

The Lévy–Itô decomposition states that

$$X_t = at + B_t + X_t^l + \lim_{\epsilon \to 0} \tilde{X}_t^\epsilon, \qquad t \geq 0, \tag{3.1}$$

where

$$X_t^l = \int_{\substack{s\in[0,t] \\ |x|>1}} x J_X(ds \times dx)$$

corresponds to discontinuous large jump process and

$$\begin{aligned} \tilde{X}_t^\epsilon &= \int_{\substack{s\in[0,t] \\ \epsilon\leq|x|\leq 1}} x\{J_X(ds \times dx) - \nu(dx)ds\} \\ &= \int_{\substack{s\in[0,t] \\ \epsilon\leq|x|\leq 1}} x\{\tilde{J}_X(ds \times dx)\} \end{aligned}$$

corresponds to compensated small jump process.

The terms in (3.1) are independent and the convergence in the last term is almost sure and uniform in t on $[0, t]$. For more details see [16] and [5].

In (3.1) the term $at + B_t$ corresponds to a continuous Gaussian Lévy process as discussed earlier. In fact every Gaussian Lévy process is continuous and can be represented in this form by two parameters: a the drift term and the variance–covariance matrix $\mathbf{A}$ of the Brownian Motion.

The term X_t^l (l symbolizing large) is a discontinuous process incorporating the jumps of X_t and is described by the Lévy measure ν. The condition $\int_{|x|\geq 1} \nu(dx) < \infty$ in (1.5) assures that X_t has a finite number of jumps having sizes larger than 1. Hence the sum

$$X_t^l = \int_{\substack{s\in[0,t] \\ |x|>1}} x\, J_X(ds \times dx) = \sum_{\substack{0\leq s\leq t \\ |\Delta X_s|\geq 1}} \Delta X_s \tag{3.2}$$

contains a.s. a finite number of terms and X_t^l is a CP process. From (2.23), with $\acute{t}_1 = 1$ and $t_1 = 0$, it follows that the jump times of this CP process is controlled by a Poisson process with rate $\nu[1, \infty)$. Also considering (2.16) we see that the distribution of jumps of this CP process is

$$dF^1(x) = \frac{\nu(dx)}{\nu[1, \infty)}\,. \tag{3.3}$$

Since $\int_{|x|\geq\epsilon} \nu(dx) < \infty$ in (1.5), then if we replace 1 in (2.23) by ϵ, where $0 < \epsilon \leq 1$, we still get a CP process:

$$X_t = \sum_{\substack{0\leq s\leq t \\ |\Delta X_s|\geq\epsilon}} \Delta X_s, \qquad t \geq 0,$$

with jump times controlled by a Poisson process with rate $\nu[\epsilon, \infty)$ and the distribution of jumps following

$$dF^\epsilon(x) = \frac{\nu(dx)}{\nu[\epsilon, \infty)}.$$

In fact it can be shown that the sum of jumps with amplitude between $\epsilon > 0$ and 1 is again a well defined CP process. But this time the number of small jumps may be infinitely many and their sum may fail to converge. To have convergence mean-subtracted jumps are summed up which leads to the compensated jump measure $\tilde{J}_X(dt \times dx) = J_X(dt \times dx) - \nu(dx)dt$ in the last term of the Lévy-Itô decomposition.

3.2 Truncated Jumps Density and the Compound Poisson Process

The process whose jumps follow a gamma distribution is known as the gamma process. From Example 1.2 it follows that this process is a Lévy process. Thus if $X_1 \sim gamma(\alpha, \beta)$ then $X_t \sim gamma(\alpha t, \beta)$. It is natural to ask if these gamma processes are also CP?

The moment generating function of X_t is thus given by

$$\mathbb{E}\Big[\exp\{sX_t\}\Big] = \left(\frac{\beta}{\beta - s}\right)^{\alpha t} = \exp\Big\{t\int_0^\infty (e^{sx}-1)\,\frac{\alpha}{x}\,e^{-\beta x}\,dx\Big\}, \qquad s < \beta, \tag{3.4}$$

where the first equality follows from Example 1.2 and the second equality is proved in Section 3.3.1.

Now the expression in (3.4) is almost of the same form as that of the moment generating function of a compound Poisson process $X_t = Z_1 + \cdots + Z_{N(t)}$ with non–negative jump sizes Z_j and common probability density function $f = f_{Z_1}$:

$$\mathbb{E}\Big[\exp\{sZ_t\}\Big] = \exp\Big\{\lambda t\int_0^\infty (e^{sx}-1)\,f(x)\,dx\Big\}, \qquad s \in \mathbb{R}. \tag{3.5}$$

Now to match the above two expressions we set:

$$\lambda\, f(x) = \frac{\alpha}{x}\,e^{-\beta x} := \lambda_0\, f^0(x), \qquad x > 0. \tag{3.6}$$

However this is not a probability density function because it is not integrable as $x \searrow 0$. To get a probability density from this, we can truncate f at $\epsilon > 0$ and write:

$$\lambda_\epsilon\, f^\epsilon(x) = \frac{\alpha}{x}\,e^{-\beta x}, \qquad x > \epsilon, \tag{3.7}$$

where f^ϵ corresponds to the density of jumps which are truncated at level $\epsilon > 0$.

Now for f^ϵ to be a probability density we just need to choose λ_ϵ appropriately. However from (3.7) we notice that $\lambda_\epsilon \to \infty$ as $\epsilon \searrow 0$. Since λ_ϵ is the rate of the Poisson process driving the compound Poisson process, we interpret this as "there will be increasingly more jumps as $\epsilon \searrow 0$".

At the same time from (3.7) we can write

$$f^\epsilon(x) = \frac{1}{\lambda_\epsilon}\,\frac{\alpha}{x}e^{-\beta x}, \qquad x > \epsilon, \tag{3.8}$$

which means that as $\epsilon \searrow 0$ the mean of the distribution with density f^ϵ, as given above, tends to zero (since $\lambda_\epsilon \to \infty$). Hence these increasingly more frequent jumps are very small.

3.3 Subordinators

Refer to the definition of finite variation of a function from Appendix-A. A Lévy process is said to be of finite variation if its trajectories are a.s. functions of finite variation.

Proposition 3.1 *A Lévy process is of finite variation if and only if its characteristic triplet* (a_c, b^2, ν) *satisfies:*

$$b^2 = 0 \qquad and \qquad \int_{|x|\le 1} \nu(dx) < \infty.$$

For a proof see [5].
Now from Remark 1.5-[5] it is known that if $\int_{|x|\le 1} |x|\, \nu(dx) < \infty$ then the corresponding cut-off function is $C(x) = 0$, which leads to the following corollary.

Corollary 3.1 *Let* $\{X_t\,;\, t \ge 0\}$ *be a Lévy process of finite variation with Lévy triplet* $(a_0, 0, \nu)$. *Then* X *can be expressed as the sum of its jumps over times* 0 *to* t *and a linear drift term:*

$$X_t = dt + \int_{[0,t]\times\mathbb{R}} x\, J_X(ds \times dx) = dt + \sum_{\substack{X_s \ne 0 \\ s\in[0,t]}} \Delta X_s\,, \qquad t \ge 0\,, \tag{3.9}$$

and its characteristic function can be expressed as

$$\mathbb{E}\big[e^{isX_t}\big] = \exp\Big[t\big\{ids + \int (e^{isx} - 1)\nu(dx)\big\}\Big] \tag{3.10}$$

where $d = a_0 - \int_{|x|\le 1} x\, \nu(dx)$.

Remark 3.1 *For a compound Poisson type Lévy process i.e. a Lévy process with all trajectories a.s. piecewise constant,* d *in* (3.10) *is zero. Another important thing to note here is that the characteristic triplet of* X *is not given by* $(d, 0, \nu)$ *but by* $(a_0, 0, \nu)$.

The variation of a Lévy process can be characterized by its Lévy measure:

$$\int_0^\infty (1 \wedge |x|)\nu(dx) = \begin{cases} \text{finite} & \text{if } X_t \text{ is of finite variation.} \\ \infty & \text{if } X_t \text{ is of unbounded variation.} \end{cases} \tag{3.11}$$

We are now in a position to define a subordinator.

Definition 3.1 *If the Gaussian component of a Lévy process $b^2 = 0$ and the Lévy measure ν is defined on $(0, \infty)$ such that $\int_0^\infty (1 \wedge x)\nu(dx) < \infty$ then the Lévy process is known as a subordinator.*

From this definition it is clear that subordinators always have positive increments and are Lévy processes with a.s. increasing trajectories.

The value of a subordinator X_t, at each $t \geq 0$, is a positive random variable. Thus subordinators are increasing (weakly) Lévy processes. It is convenient to describe them in terms of their Laplace transform (since here it always exists) instead of their Fourier transform. Consequently (3.10) implies that

$$\mathbb{E}\big[e^{sX_t}\big] = \exp\Big[t\big\{ds + \int (e^{sx} - 1)\,\nu(dx)\big\}\Big] = e^{tl(-s)}, \qquad t \geq 0\,,\ s \in \mathbb{R}\,, \tag{3.12}$$

where $l(s)$ where is called the Laplace exponent of the subordinator X_t.

All these facts yield the following proposition which completely characterizes a subordinator.

Proposition 3.2 *Let $\{X_t\,;\, t \geq 0\}$ be a Lévy process on $\mathbb{R}$. The following conditions are equivalent:*

[i] $X_t \geq 0$ a.s. for some $t > 0$.

[ii] $X_t \geq 0$ a.s. for every $t > 0$.

[iii] Sample paths of X_t are almost surely non-decreasing:

$$t \geq \acute{t} \qquad \Rightarrow \qquad X_t \geq X_{\acute{t}} \quad a.s.$$

[iv] The characteristic triplet of X_t satisfies $b^2 = 0$, $\nu(-\infty, 0] = 0$, $\int_0^\infty (1 \wedge x)\nu(dx) < \infty$ and $d = a_0 - \int_{|x|\leq 1} x\nu(dx) \geq 0$, that is X_t has no diffusion component, only positive jumps of finite variation and positive drift.

For a proof see [5].

We end this section with a nice construction showing how subordinators can be obtained from Poisson point process.

Theorem 3.1 *Let $d \geq 0$ and let $(\triangle_t)_{t\geq 0}$ be a Poisson point process with intensity function $g : (0,\infty) \to [0,\infty)$ (or intensity measure ν on $(0,\infty)$) such that*

$$\int_0^\infty (1 \wedge x)\, g(x)\, dx < \infty \qquad \left(or \int_{(0,\infty)} (1 \wedge x)\, \nu(dx) < \infty \right),$$

then the process $X_t = \{dt + \sum_{t\leq t} \triangle_t\, ;\, t \geq 0\}$ is a subordinator with moment generating function $\mathbb{E}\Big(\exp\{sX_t\}\Big) = \exp\{t\Psi(s)\}$, where

$$\Psi(s) = ds + \int_0^\infty (e^{sx} - 1)\, g(x)\, dx \qquad \left(or\ \ \Psi(s) = ds + \int_{(0,\infty)} (e^{sx} - 1)\, \nu(dx) \right).$$

For a proof of see [19].

Note that the g in Theorem 3.1, in (2.24) and (2.25) is the same.

3.3.1 Examples of subordinators

Here we will discuss some frequently used subordinators. We will derive in closed form the Laplace exponents of gamma and α-stable subordinators hence getting their moment generating functions.

Poisson Subordinator

Comparing (2.11) with (3.12) we can easily see that the Poisson process $\{N_t\, ;\, t \geq 0\}$ is a subordinator with $d = 0$ and $\nu(dx) = \lambda\, \delta_1(x)dx$ where δ_1 is the discrete unit point mass at jump size 1.

Compound Poisson Subordinator

Comparing (2.13) with (3.12) we can conclude that the CP process is a subordinator with $d = 0$ and Lévy measure $\nu(dx) = \lambda f(x)\, dx$, for $x > 0$, where f is the probability density function of the independent and identically distributed non-negative jump sizes.

Gamma Subordinator

The process with gamma distributed increments is known as the gamma process. From Example 1.2 it follows that this process is a Lévy process. Thus if $X_1 \sim gamma(\alpha, \beta)$ then

$X_t \sim gamma(\alpha t, \beta)$. The moment generating function and Laplace exponent are given by:

$$\mathbb{E}\big(\exp\{sX_t\}\big) = \left(\frac{\beta}{\beta - s}\right)^{\alpha t} = \exp\Big\{t\int_0^\infty (e^{sx}-1)\,\frac{\alpha}{x}\,e^{-\beta x}\,dx\Big\} = e^{tl(-s)}\,, \qquad s < \beta\,. \quad (3.13)$$

Hence by (3.12), $X = \big\{X_t\,;\,t \geq 0\big\}$ is a subordinator with $d = 0$ and the Lévy measure defined as $\nu(dx) = \frac{\alpha}{x}e^{-\beta x}\,dx$. Its Laplace exponent is given by:

$$l(s) = \int_0^\infty (e^{-sx} - 1)\frac{\alpha}{x}e^{-\beta x}dx, \qquad s < \beta. \tag{3.14}$$

In (3.13) the first equality follows from Example 1.2, however the integral in the second equality is mathematically tricky to evaluate. Here are the details.

For simplicity let $\alpha = 1$ and assume that

$$\int_0^\infty (1 - e^{-sx})\frac{1}{x}e^{-\beta x}dx = g(s), \qquad \text{with } g(0) = 0.$$

Observe that for fixed $\beta > 0$ both sides equal 0 at $s = 0$. It is therefore sufficient to show that g is such that the $s-$derivative on both sides coincides.

That is

$$\frac{d}{ds}\left(\int_0^\infty (1 - e^{-sx})\frac{1}{x}e^{-\beta x}dx\right) = g'(s). \tag{3.15}$$

Now to check that on left side the differentiation under the integral is plausible we need to ensure that after differentiating the integrand the result will not diverge. This is true as:

$$\frac{d}{ds}[(1 - e^{-sx})\frac{1}{x}e^{-\beta x}] = e^{-sx}e^{-\beta x} \leq e^{-\beta x},$$

where $\beta > 0$ can be fixed in such a way that the dominance holds for any $s > 0$. Figure 3.1 shows a typical example of such a dominance.

Since $x \mapsto e^{-\beta x}$ is integrable on $[0, \infty)$, for $\beta > 0$, the differentiation under the integral is plausible. Hence we have:

$$\int_0^\infty e^{-sx}e^{-\beta x}dx = g'(s).$$

That is,

$$\frac{1}{\beta + s} = g'(s),$$

and hence

$$g(s) = \int_0^s \frac{1}{\beta + z}\,dz + C = \log\big(\frac{\beta + s}{\beta}\big)\,,$$

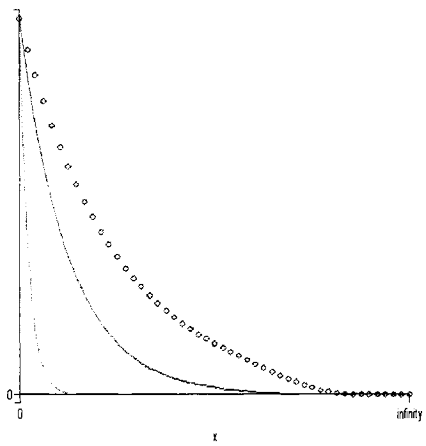


Figure 3.1: Dominance of $e^{-\beta x}$ (dotted–line) over $e^{-sx}e^{-\beta x}$ for $\beta = \min\{s \mid s = 0.5$ (red), $s = 5$ (green), $s = 50$ (yellow)$\}$.

where the constant C is assumed equal to 0 without loss of generality. Hence

$$-l(s) = \alpha \log\big(\frac{\beta + s}{\beta}\big) \qquad \text{and so} \qquad l(-s) = \log\big(\frac{\beta}{\beta - s}\big)^{\alpha}, \qquad s < \beta.$$

Finally

$$e^{tl(-s)} = e^{\log(\frac{\beta}{\beta - s})^{\alpha t}} = \big(\frac{\beta}{\beta - s}\big)^{\alpha t}, \qquad t \geq 0 \text{ and } s < \beta.$$

Hence the second equality in (3.13) holds.

α-stable Subordinator

The process with Lévy-Khinchine representation given by:

$$\mathbb{E}\Big[\exp\{-sX_t\}\Big] = \exp\big\{t\int_0^{\infty}(e^{-sx} - 1)\frac{1}{x^{\alpha+1}}dx\big\}, \qquad s > 0, \tag{3.16}$$

is known as the α-stable Lévy process. Comparing this expression with (3.12) we see that $X = \{X_t\,;\, t \geq 0\}$ is a subordinator with $d = 0$ and $\nu(dx) = \frac{1}{x^{\alpha+1}}dx$. Its Laplace exponent is given by:

$$l(s) = \int_0^{\infty}(e^{-sx} - 1)\frac{1}{x^{\alpha+1}}dx, \qquad s > 0. \tag{3.17}$$

Like in the previous example, obtaining the closed form

$$\exp\big\{t\int_0^{\infty}(e^{-sx} - 1)\frac{1}{x^{\alpha+1}}dx\big\} = \exp\big\{t\frac{\Gamma(1-\alpha)}{\alpha}s^{\alpha}\big\}, \qquad s > 0, \tag{3.18}$$

is mathematically tricky. As before we are looking for h satisfying:

$$\int_0^\infty (e^{-sx}-1)\frac{1}{x^{\alpha+1}}dx = h(s), \qquad \text{with} \qquad h(0)=0,$$

and for fixed $\alpha > 0$

$$\frac{d}{ds}\left(\int_0^\infty (1-e^{-sx})\frac{1}{x^{\alpha+1}}dx\right) = h'(s), \tag{3.19}$$

up to a multiplicative constant. Differentiation under the integral is plausible since:

$$\frac{d}{ds}\left((1-e^{-sx})\frac{1}{x^{\alpha+1}}\right) = \frac{e^{-sx}}{x^\alpha} \leq \frac{1}{x^\alpha}, \quad \text{for a suitably fixed } \alpha > 0,$$

where $x \mapsto \frac{1}{x^\alpha}$, for $\alpha > 0$, is integrable on $[0,\infty)$. A typical example of such dominance appears in Figure 3.2. Hence we have

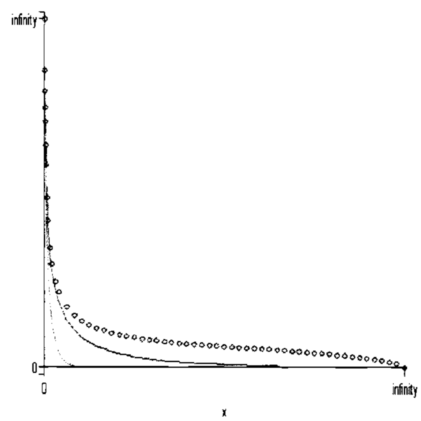


Figure 3.2: Dominance of the graph of $\frac{1}{x^\alpha}$ (dotted-line) over the graphs of $\frac{e^{-sx}}{x^\alpha}$ for $\alpha = \beta = \min\{s \mid s = 0.5 \text{ (red)}, s = 5 \text{ (green)}, s = 50 \text{ (yellow)}\}$.

$$\int_0^\infty x^{(1-\alpha)-1}e^{-sx}dx = h'(s) \quad \Rightarrow \quad \frac{\Gamma(1-\alpha)}{s^{(1-\alpha)}} = h'.(s)$$

That is

$$h(s) = \int_0^s \frac{\Gamma(1-\alpha)}{z^{(1-\alpha)}}dz = \frac{\Gamma(1-\alpha)}{\alpha}\{s^\alpha - 0\} = \frac{\Gamma(1-\alpha)}{\alpha}s^\alpha, \qquad s > 0.$$

Therefore $l(s) = \frac{\Gamma(1-\alpha)}{\alpha}s^\alpha$ and hence for $s > 0$ and $t \geq 0$

$$\mathbb{E}(\exp\{-sX_t\}) = \exp\{t\int_0^\infty (e^{-sx}-1)\frac{1}{x^{\alpha+1}}dx\} = \exp\{t\frac{\Gamma(1-\alpha)}{\alpha}(s)^\alpha\} \ . \tag{3.20}$$

Remark 3.2 [On the α-stable subordinator]

[1] From the expression in (3.20) we have the notion of stability:

$$\mathbb{E}\big[\exp\{-sc^{\frac{1}{\alpha}}X_{\frac{t}{c}}\}\big] = \exp\{\frac{t}{c}\frac{\Gamma(1-\alpha)}{\alpha}(sc^{\frac{1}{\alpha}})^{\alpha}\} = \exp\{t\frac{\Gamma(1-\alpha)}{\alpha}s^{\alpha}\},$$

which implies that $c^{\frac{1}{\alpha}}X_{\frac{t}{c}} \sim X_t$.

[2] More generally one can consider the tempered stable processes with $\nu(dx) = \frac{1}{x^{\alpha+1}}\exp\{-\rho x\}dx$ with the tempering factor $\rho > 0$.

[3] For an α-stable process to be a subordinator we must have α between 0 and 1.

Inverse Gaussian Subordinator

Sparing details we mention that equation (1.2) can be written in the form of equation (3.12). Then it can be shown that if $X_1 \sim IG(\mu, \theta)$, the inverse Gaussian process $X_t \sim IG(\mu t, \theta)$ is a subordinator with Lévy measure $\nu(dx) = \frac{\mu}{\sqrt{2\pi}x^{3/2}}\exp\big\{-\frac{1}{2}\theta^2 x\big\}dx$, $x > 0$.

The next chapter focuses on the estimation problem of subordinator for the aggregate claims process in insurance settings. A recent work on parametric estimation of subordinator is [12].

Chapter 4

Estimation of ν and λ

We saw, in the previous chapter, that the estimation of ν is of great importance in revealing the distribution of jumps of size greater than some threshold $\epsilon > 0$. This distribution corresponds to F^{ϵ}, the distribution of claims larger than the threshold ϵ in an increasing insurance claims process driven by a subordinator. In this chapter we discuss the estimation of ν as well as that of λ for the underlying subordinator.

4.1 Insurance Aggregate Claims Process

Consider the aggregate claims process of an insurance company. A close look reveals that this can be modelled as a Lévy process, or more precisely as a subordinator. Here the positive drift may be generated by small claims. This liability may not be a problem for the company as its impact may be accommodated by redistributing different assets. But the insurance companies remain cautious about the occurrence of claims larger than some threshold. This threshold itself is a parameter to be determined by the insurance business. So knowledge about the distribution and rate of claims larger than a certain threshold is of interest in an insurance business.

Here our goal is to estimate the rate λ_{ϵ} and the distribution F^{ϵ} of claims (or jumps) larger than some threshold $\epsilon > 0$ from a set of observed values sampled from the claims subordinator process. There are many other exciting applications of subordinators in finance and insurance. For a recent application of a subordinator in Risk theory we refer to [8]. In the case when the process is of jump-diffusion type, rather than a subordinator,

[11] is a recent reference with applications to both in insurance and finance.

There may be three different sampling schemes:

Scheme 1: Estimation of a Lévy driven increasing jump process based on continuous time realizations such as $\{X_s\,;\, 0 \le s \le t\}$.

Scheme 2: Estimation based on a set of discrete time observations, say, $X_{t_1}, X_{t_2}, \cdots, X_{t_n}$ at fixed time points $t_1, t_2, \cdots, t_n$, of a Lévy driven increasing jump process.

Scheme 3: Estimation based on a set of discrete random time observations, say $X_{T_1}, \cdots, X_{T_n}$, where the random time points $T_1, T_2, \cdots$ represents claim arrival times.

With these observations it is of interest to estimate λ and ν which in turn reveals the distribution of jumps of a given magnitude falling in some Borel set $A \subset [0, \infty)$.

4.2 Some Motivations

We discuss an estimation procedure for the jump function associated with a Lévy process. In Theorem 2.1 we saw that the distribution of X_t is infinitely divisible for each t. The logarithm of the characteristic function of X_t, can be written as:

$$\log \Phi_{X_t}(s) = iats + t \int_{-\infty}^{\infty} \left[e^{isx} - 1 - \frac{isx}{(1+x^2)} \right] \frac{(1+x^2)}{x^2}\, dG(x)\,, \qquad s \in \mathbb{R}\,. \tag{4.1}$$

This is known as the canonical or Kolmogorov representation, see [13], [10]. Here a is a real constant and G is a bounded non–decreasing function with $G(-\infty) = 0$. This G is known as the jump function. Section 4.3.3 shows how this representation of the logarithm of the characteristic function is related to the logarithm of the characteristic function appearing in (2.4). We recall that the distribution of X_t is the convolution of a normal law and a possibly infinite number of Poisson laws. The followings are two important properties.

Property 4.1 *The function G in* (4.1) *satisfies the following properties:*

[1] The jump of G at $x = 0$, i.e. $\sigma^2 = G(0^+) - G(0^-)$, is the variance of the normal component.

[2] Its set of points of increase for $x \neq 0$ gives information as to the nature of the Poisson components, i.e. the relative density of the magnitude of the discontinuities of the sample function.

We explore these facts further, since we believe these are the motivating factors in the application of the theory of infinite divisibility and hence the theory of Lévy processes. First let us discuss the relationship between the jump function G and infinite divisibility.

4.2.1 Jump function G and infinite divisibility

To make things simpler let us consider the following lemma.

Lemma 4.1 *In the case where the random variable X_t in (4.1) has a finite variance:*

[1] There exists a real constant a^* and a bounded non-decreasing function H such that (4.1) takes the form

$$\log \Phi_{X_1}(s) = ia^*s + \int_{-\infty}^{\infty} (e^{isx} - 1 - isx)\frac{1}{x^2} dH(x), \qquad s \in \mathbb{R}. \tag{4.2}$$

[2] The random variable X_t has finite variance if and only if

$$\int_{-\infty}^{\infty} (1 + x^2) dG(x) < \infty, \tag{4.3}$$

and then the relation of H to G becomes

$$H(s) = \int_{-\infty}^{s} (1 + x^2) dG(x), \tag{4.4}$$

for all $s \in C(H)$ and that of a to a^* is

$$a^* = a + \int_{-\infty}^{\infty} x dG(x). \tag{4.5}$$

Proof. See [10] for Part-[1], which is in fact the proof of (4.1) for a simpler case. A rigorous treatment of the relations in Part-[2] can be found in [13]. □

Here we recall Remark (1.5)-[5], suggesting the cut-off function $C(x) = 1$ when finiteness of the variance of the distribution implies finiteness of the Lévy measure. For more general results on the finiteness of the integrals involving Lévy processes we refer to [6].

The following proposition characterizes the relationship of jump function G with infinitely divisible distributions.

Proposition 4.1 *Equation* (4.2) *is the characteristic function of an infinitely divisible random variable* X_1 *such that*

$$\mathbb{E}(X_1) = a^* \qquad \text{and} \qquad \mathbb{V}(X_1) = \int_{\mathbb{R}} dH(x) = b^*.$$

Proof. Differentiating (4.2) with respect to s we get:

$$\Phi'_{X_1}(s) = e^{\left\{ia^*s+\int_{-\infty}^{\infty}(e^{isx}-1-isx)\frac{1}{x^2}dH(x)\right\}}\left[ia^* + \int_{-\infty}^{\infty}(ixe^{isx} - ix)\frac{1}{x^2}dH(x)\right]. \tag{4.6}$$

So $\Phi'_{X_1}(0) = e^0[ia^* + 0]$. Thus $i\mathbb{E}(X_1) = \Phi'_{X_1}(0)$ implies that $\mathbb{E}(X_1) = a^*$.
Differentiating (4.6) with respect to s we get:

$$\Phi''_{X_1}(s) = e^{\left\{ia^*s+\int_{-\infty}^{\infty}(e^{isx}-1-isx)\frac{1}{x^2}dH(x)\right\}}\left[ia^* + \int_{-\infty}^{\infty}(ixe^{isx} - ix)\frac{1}{x^2}dH(x)\right]^2 \tag{4.7}$$

$$+e^{\left\{ia^*s+\int_{-\infty}^{\infty}(e^{isx}-1-isx)\frac{1}{x^2}dH(x)\right\}}\left[\int_{-\infty}^{\infty}(i^2x^2e^{isx})\frac{1}{x^2}dH(x)\right],$$

so

$$\Phi''_{X_1}(0) = e^0[ia^*]^2 - e^0[\int_{\infty}^{\infty} dH(x)] = -a^{*2} - b^*.$$

Then $i^2\mathbb{E}(X_1^2) = \Phi''_{X_1}(0) = -a^{*2} - b^*$ implies that $\mathbb{E}(X_1^2) = a^{*2} + b^*$. Hence $\mathbb{V}(X_1) = b^*$. □

Remark 4.1 *Considering* (4.3), (4.4) *and* (4.5), *a similar result can be proved for* (4.1).

4.2.2 Explanation of Property 4.1

Let us come back to Property 4.1. Assume that H has a point mass of σ^2 (it will be clear later why we use the symbol σ^2) at $x = 0$ and a density of λx^2 at $x \neq 0$. Then from (4.2) the characteristic function of the process in (4.1) can be written as:

$$\log \Phi_{X_1}(s) = ia^*s + \int_{\{0\}}(e^{isx} - 1 - isx)\frac{1}{x^2}dH(x) + \int_{\mathbb{R}\setminus\{0\}}(e^{isx} - 1 - isx)\frac{1}{x^2}dH(x). \tag{4.8}$$

To see what is the value of the first integrand as $x \to 0$, consider

$$\begin{aligned}\lim_{x\to 0}(e^{isx} - 1 - isx)\frac{1}{x^2} &= \lim_{x\to 0}\frac{ise^{isx} - is}{2x}, && \text{by L'Hopitals rule,}\\ &= \lim_{x\to 0}\frac{(is)^2e^{isx}}{2} = \frac{-s^2}{2}. && (4.9)\end{aligned}$$

Then (4.8) becomes:

$$\Phi_{X_1}(s) = \exp\left\{ia^*s - \frac{s^2}{2}\int_{\{0\}}dH(x)\right\}\exp\left\{(e^{isx} - 1 - isx)\lambda\right\}, \qquad s \in \mathbb{R}, \tag{4.10}$$

$$= \underbrace{\exp\left\{ia^*s - \frac{s^2}{2}[H(0^+) - H(0^-)]\right\}}_{Part-I} \underbrace{\exp\left\{\lambda(e^{isx} - 1)\right\}}_{Part-II} \underbrace{\exp\left\{is(-x\lambda)\right\}}_{Part-III}, \tag{4.11}$$

where Part-I is clearly the characteristic function of say $Y_1 \sim N(a^*, \tilde{b})$ and $\tilde{b} = \int_{\{0\}} dH(x) = H(0^+) - H(0^-)$. This explains why we represent the point mass at $x = 0$ with the symbol σ^2 and proves the first claim that the jump $H(0^+) - H(0^-)$ of H at $x = 0$ is the variance of the normal component. The same is true for G since from (4.4), with condition (4.3) guaranteed by the finite variance, we have:

$$\frac{dH(s)}{ds} = (1 + s^2)\frac{dG(s)}{ds},$$

which implies that

$$H(0^+) - H(0^-) = dH(0) = dG(0) = G(0^+) - G(0^-).$$

From (4.5) at $x = 0$ we have $a = a^*$.

Part-II is the characteristic function of a $Poisson(\lambda)$, but not at s, rather at sx. Since we know that for any random variable X,

$$\Phi_X(ct) = \int e^{i(ct)x} dF_X(x) = \int e^{it(cx)} dF_X(x) = \Phi_{cX}(t), \qquad t \in \mathbb{R}.$$

Hence Part-II represents the characteristic function:

$$\Phi_{Pois(\lambda)}(sx) = \Phi_{xPois(\lambda)}(s), \qquad s \in \mathbb{R},$$

of a random variable, say $Y_2 = xPois(\lambda)$.

Finally Part-III is clearly the characteristic function of $\mathbb{I}_{\{Y_3 = -x\lambda\}}$. Hence $\Phi_{X_1}(s)$ is the product of three characteristic functions $\Phi_{Y_1}(s)$, $\Phi_{Y_2}(s)$ and $\Phi_{Y_3}(s)$. Now by the uniqueness of characteristic function we can conclude that $X_1 = Y_1 + Y_2 + Y_3$ where Y_i's are independent. Therefore $F_{X_1} = F_{Y_1} * F_{Y_2} * F_{Y_3}$.

Now we can easily see that

$$\mathbb{E}(X_1) = \mathbb{E}(Y_1 + Y_2 + Y_3) = a^* + x\lambda - x\lambda = a^*,$$

and

$$\mathbb{V}(X_1) = \mathbb{V}(Y_1 + Y_2 + Y_3) = \sigma^2 + x^2\lambda + 0 = b^*.$$

This explains Proposition 4.1.

Again from (4.8) we can write:

$$\Phi_{X_1}(s) = \Phi_{N(a^*,\sigma^2)}(s)\ \exp\{\int_{\mathbb{R}\setminus\{0\}} (e^{isx} - 1 - isx)\frac{1}{x^2}dH(x)\}, \qquad s \in \mathbb{R}, \tag{4.12}$$

where we have seen that $\exp\{\lambda(e^{isx} - 1 - isx)\}$ is the characteristic function of $Y_2 - x\,\lambda = x\big[Pois(\lambda) - \lambda\big]$, a transformed centered Poisson random variable. Since $e^{\lambda(e^{isx}-1-isx)} = \left[e^{\frac{\lambda}{n}(e^{isx}-1-isx)}\right]^n$, (4.12) implies that:

$$\Phi_{X_1}(s) = \Phi_{N(a^*,\sigma^2)}(s)\left[\Phi_{\{x(Pois(\frac{\lambda}{n})-\frac{\lambda}{n})\}}(s)\right]^n.$$

It means that X_1 is the convolution of a $N(a^*, \sigma^2)$ variable and the n-fold convolution of a $x\big[Pois(\lambda) - \lambda\big]$ random variable.

In general:

$$\begin{aligned}\Phi_{X_1}(s) &= \Phi_{N(a^*,\sigma^2)}(s)\ \exp\left\{\int_{\mathbb{R}\setminus\{0\}} (e^{isx} - 1 - isx)\frac{1}{x^2}dH(x)\right\} \\ &= \Phi_{N(a^*,\sigma^2)}(s)\ \exp\left\{\sum_{i=-\infty}^{\infty}\int_{a_i}^{a_{i+1}} (e^{isx} - 1 - isx)\frac{1}{x^2}dH_i(x)\right\}\end{aligned} \tag{4.13}$$

where $\bigcup_{i=-\infty}^{\infty}[a_i, a_{i+1}] = \mathbb{R}\setminus\{0\}$, $H_i(x) = H(x)\mathbb{I}_{[a_i,a_{i+1}]}(x)$ and $H(a_{i+1}) - H(a_i) \neq 0$. This last equation summarizes the spirit of Property 4.1-[2]. It says that H_i is the relative density of the jumps of size in the interval $[a_i, a_{i+1}]$.

4.3 Estimating G

Assume that the experimenter can observe a sample function of X at any finite number of chosen values of t. Accordingly for any integer n, let

$$Y_k^n = X_{nk} \overset{D}{=} X_{\frac{k}{n}} - X_{\frac{(k-1)}{n}}\,, \qquad k = 1, \ldots, nN\,.$$

Then the following estimator of G in (4.1),

$$G^*_{N,n}(s) = \frac{1}{N}\sum_{1}^{nN}\frac{(Y_k^n)^2}{\big[1 + (Y_k^n)^2\big]}\mathbb{I}_{[Y_k^n \le s]},$$

which is proposed in [15], is proved to be strongly consistent in the following sense:

$$\mathbf{P}\big\{\lim_{n\to\infty}\lim_{N\to\infty} G^*_{N,n}(s) = G(s) \quad \text{for all} \quad s \in C(G)\big\} = 1,$$

where $C(G)$ denotes the set of all values of s at which G is continuous. Here $N = [Tn]$, T being the largest observed time value. One drawback of this estimate is that it is not necessarily an unbiased estimate of $G(s)$, at all $s \in C(G)$.

If we have an estimator of G, we can then relate it to that of the Lévy measure ν.

4.3.1 Description of the method for estimating G

The method is based solely on necessary conditions for a general central limit theorem for sums of independent random variables. The statement of the theorem can be found in [9], and is reproduced here as follows.

Theorem 4.1 *In order for some suitably chosen constants $A_1, A_2, \cdots, A_n$ and the distribution function of sums*

$$W_n = Z_{n,1} + Z_{n,2} + \cdots + Z_{n,k_n} - A_n$$

of independent random variables $Z_{n,k}$, with distribution $F_{n,k}$, to converge to a limiting distribution function, it is necessary and sufficient that there exist a bounded non–decreasing function G such that

$$\sum_{k=1}^{k_n} \int_{-\infty}^{s} \frac{z^2}{1+z^2} dF_{n,k}(z+\beta_{n,k}) \to G(s), \qquad s \in C(G), \tag{4.14}$$

as $n \to \infty$ and where $\beta_{n,k} = \int_{-r}^{r} z dF_{n,k}(z)$ for arbitrary $r > 0$.

One consequence of this theorem is that A_n can be written as a sum $A_n = \beta_{n,1} + \beta_{n,2} + \cdots + \beta_{n,k_n}$. Hence

$$W_n = (Z_{n,1} - \beta_{n,1}) + (Z_{n,2} - \beta_{n,2}) + \cdots + (Z_{n,k_n} - \beta_{n,k_n}).$$

Now since the limit of infinitely divisible distributions is infinitely divisible, then the logarithm of the characteristic function of the limiting distribution is of the form

$$\log \Phi(s) = ias + \int_{-\infty}^{\infty} \left(e^{isz} - 1 - \frac{isz}{1+z^2}\right) \frac{(1+z^2)}{z^2} dG(z)\,, \qquad s \in \mathbb{R}\,. \tag{4.15}$$

The function G in (4.15) is the same as in (4.14) which leads to avail an estimator of G. By necessity and sufficiency in Theorem 4.1 the pair (a, G) determines and is determined by this limiting distribution.

In case the limiting distribution has a finite variance it is known that its characteristic function is of the form (4.2) and then the relation of H to G is given by (4.4).

In the case of the Lévy process $X = \{X_t; t \geq 0\}$ we have

$$Y_k^n = X_{\frac{k}{n}} - X_{\frac{k-1}{n}}, \qquad \text{for } k = 1, 2, \cdots, nN.$$

Letting $n \in \mathbb{N}^+$ we see that our infinitesimal system of random variables satisfies

$$\begin{aligned} Y_1^n + Y_2^n + \cdots + Y_n^n &= (X_{\frac{1}{n}} - X_0) + (X_{\frac{2}{n}} - X_{\frac{1}{n}}) + \cdots + (X_1 - X_{\frac{n-1}{n}}) \\ &= X_1. \end{aligned}$$

Hence the distribution of X_1 is the limit law of the distributions of this sequence of sums. In this case since $Y_k^n = X_{\frac{k}{n}} - X_{\frac{k-1}{n}} \stackrel{D}{=} X_{\frac{1}{n}}$ (which is independent of k) and $F_{n,k}(y)$ is, for fixed n, the same for all k (n fixed), hence it is denoted as $F_n(y)$. Consequently $\beta_{n,k} = \beta_n = \int_{-\tau}^{\tau} y dF_n(y)$ for $1 \leq k \leq n$. Then using (4.14) we have

$$\int_{-\infty}^{s} \frac{y^2}{(1+y^2)} dF_n(y+\beta_n) + \cdots + \int_{-\infty}^{s} \frac{y^2}{(1+y^2)} dF_n(y+\beta_n) \longrightarrow G(s), \tag{4.16}$$

that is

$$n \int_{-\infty}^{s} \frac{y^2}{(1+y^2)} dF_n(y+\beta_n) \longrightarrow G(s), \qquad s \in C(G)$$

as $n \to \infty$. Thus with

$$G_n(s) = n \int_{-\infty}^{s} \frac{y^2}{(1+y^2)} dF_n(y+\beta_n), \qquad s \in \mathbb{R}, \tag{4.17}$$

we have $G_n(s) \to G(s)$ for all $s \in C(G)$. So clearly the presence of β_n is the only problem in establishing an estimator for $G(s)$.

Now we show that for our purpose the presence of β_n is redundant, i.e. letting

$$\tilde{G}_n(s) = n \int_{-\infty}^{s} \frac{y^2}{(1+y^2)} dF_n(y), \qquad s \in \mathbb{R}, \tag{4.18}$$

we need to show that $\tilde{G}_n(s) \to G(s)$, as $n \to \infty$, for all $s \in C(G)$.

Accordingly, let

$$G_n^{**}(s) = n \int_{-\infty}^{s} \frac{(y-\beta_n)^2}{[1+(y-\beta_n)^2]} dF_n(y), \qquad s \in \mathbb{R}. \tag{4.19}$$

Using (4.17) and (4.19), we get a relation between $G_n(s)$ and $G_n^{**}(s)$ as:

$$\begin{aligned} G_n^{**}(s+\beta_n) &= n \int_{-\infty}^{s+\beta_n} \frac{(y-\beta_n)^2}{[1+(y-\beta_n)^2]} dF_n(y) \\ &= n \int_{-\infty}^{s} \frac{y^2}{(1+y^2)} dF_n(y+\beta_n) = G_n(s), \qquad s \in \mathbb{R}. \end{aligned} \tag{4.20}$$

We now need to prove the following results. These results can be found in [15] but we reproduce them in a more accessible form. The first result is on the convergence of $G_n^{**}(s)$ defined in (4.19).

Lemma 4.2 *$G_n^{**}(s) \to G(s)$ as $n \to \infty$, for all $s \in C(G)$.*

Proof. For $s \in C(G)$ fixed and arbitrary $\epsilon > 0$, there exist a $\delta > 0$ and an integer N, such that $s \pm \delta \in C(G)$ satisfy

$$G(s+\delta) - G(s-\delta) < \frac{\epsilon}{2} \qquad \text{and} \qquad G_n(s+\delta) - G_n(s-\delta) < \epsilon,$$

while $|\beta_n| < \delta$ for all $n > N$. The inequality involving G_n follows because we have $G_n(s) \to G(s)$ implying that

$$\Big[G_n(s+\delta) - G_n(s-\delta)\Big] \in \big(G(s) - \epsilon, G(s) + \epsilon\big).$$

The above inequalities imply that

$$\Big|\big[G_n(s+\delta) - G_n(s-\delta)\big] - \big[G(s+\delta) - G(s-\delta)\big]\Big| < \frac{\epsilon}{2}. \tag{4.21}$$

Hence for all $n > N$,

$$\begin{aligned}
\big|G_n^{**}(s+\beta_n) - G_n^{**}(s)\big| &= \Big|n\int_s^{s+\beta_n} \frac{(y-\beta_n)^2}{[1+(y-\beta_n)^2]} dF_n(y)\Big| \\
&= \big|G_n(s) - G_n(s-\beta_n)\big|, \qquad \text{by (4.20)} \\
&\le \big|G_n(s+\delta) - G_n(s-\delta)\big| \\
&\le \big|[G_n(s+\delta) - G_n(s-\delta)] - [G(s+\delta) - G(s-\delta)]\big| \\
&\qquad + \big|G(s+\delta) - G(s-\delta)\big|, \qquad \text{using (4.21)}, \\
&< \epsilon,
\end{aligned}$$

completing the proof. □

With the convergence of $G_n^{**}(s)$ we now prove the convergence of $\tilde{G}_n(s)$.

Lemma 4.3 *If $s < 0$ and if $s \in C(G)$, then $\tilde{G}_n(s) \to G(s)$ as $n \to \infty$.*

Proof. In order to prove this lemma we consider the function

$$f_n(y) = \frac{y^2}{(1+y^2)} \frac{[1+(y-\beta_n)^2]}{(y-\beta_n)^2}, \qquad y \le s, \tag{4.22}$$

where $s \in C(G)$ is fixed. Then $f_n(y)$ is finite for all sufficiently large n.

Now from (4.19) we have

$$\frac{dG_n^{**}(s)}{dF_n(s)} = n\frac{(s-\beta_n)^2}{[1+(s-\beta_n)^2]},$$

that is

$$dG_n^{**}(y) = n\frac{(y-\beta_n)^2}{[1+(y-\beta_n)^2]}dF_n(y).$$

Therefore, from (4.22) and (4.18)

$$\int_{-\infty}^{s} f_n(y)dG_n^{**}(y) = \tilde{G}_n(s). \tag{4.23}$$

Hence

$$\begin{aligned} |\tilde{G}_n(s) - G(s)| &\leq \left|\int_{-\infty}^{s} f_n(y)dG_n^{**}(y) - \int_{-\infty}^{s} dG_n^{**}(y)\right| + \left|\int_{-\infty}^{s} dG_n^{**}(y) - G(s)\right| \\ &\leq \sup_{y\leq s}\left|f_n(y) - 1\right|dG_n^{**}(s) + \left|G_n^{**}(s) - G(s)\right| \longrightarrow 0 \text{ as } n\to\infty, \end{aligned}$$

where the last convergence follows from Lemma 4.2 and the fact that $f_n(y)$ converges uniformly to 1 over any closed set not containing zero. Hence the proof is complete. □

With a similar argument the following result holds (see [15] for a proof).

Lemma 4.4 *If $0 < a < b$ and if $a, b \in C(G)$, then*

$$\{\tilde{G}_n(b) - \tilde{G}_n(a)\} \longrightarrow \{G(b) - G(a)\},$$

as $n \to \infty$.

Lemma 4.2 and 4.3 can now be used to prove the following more general theorem on convergence.

Theorem 4.2 *If $s \in C(G)$, then $\tilde{G}_n(s) \to G(s)$.*

Proof. Since in the case when $s < 0$ the convergence is shown in Lemma 4.3, we need to prove the theorem only in the case when $s > 0$. We begin by establishing two inequalities which are the building blocks for the proof:

$$\frac{1}{[1+(\tau+|\beta_n|)^2]}\left\{n\int_{-\tau}^{\tau} y^2 dF_n(y) - 2n\beta_n^2 + n\beta_n^2\big[F_n(\tau) - F_n(-\tau)\big]\right\}$$

$$
\begin{aligned}
&= \frac{1}{[1+(\tau+|\beta_n|)^2]} n \int_{-\tau}^{\tau} (y-\beta_n)^2 dF_n(y), \qquad \text{for any } \tau \in \mathbb{R}, \\
&\le n \int_{-\tau}^{\tau} \frac{(y-\beta_n)^2}{[1+(y-\beta_n)^2]} dF_n(y) \le n \int_{-\tau}^{\tau} (y-\beta_n)^2 dF_n(y) \qquad (4.24) \\
&= n \int_{-\tau}^{\tau} y^2 dF_n(y) - 2n\beta_n \int_{-\tau}^{\tau} y dF_n(y) + n\beta_n^2 \big[F_n(\tau) - F_n(-\tau)\big] \\
&\le n \int_{-\tau}^{\tau} y^2 dF_n(y) - n\beta_n^2, \quad \text{since } \big[F_n(\tau) - F_n(-\tau)\big] \le 1, \qquad (4.25)
\end{aligned}
$$

and

$$
n \int_{-\tau}^{\tau} y^2 dF_n(y) \le (1+\tau^2) n \int_{-\tau}^{\tau} \frac{y^2}{(1+y^2)} dF_n(y) \le (1+\tau^2) n \int_{-\tau}^{\tau} y^2 dF_n(y). \qquad (4.26)
$$

Now with the above inequalities we obtain:

$$
\begin{aligned}
n \int_{-\tau}^{\tau} \frac{y^2}{(1+y^2)} dF_n(y) &\le n \int_{-\tau}^{\tau} y^2 dF_n(y) \\
&\le [1+(\tau+|\beta_n|)^2] n \int_{-\tau}^{\tau} \frac{(y-\beta_n)^2}{[1+(y-\beta_n)^2]} dF_n(y) \qquad (4.27) \\
&\qquad + [2 - F_n(\tau) + F_n(-\tau)] n\beta_n^2, \qquad \text{by (4.24)}, \\
&\le [1+(\tau+|\beta_n|)^2] n \int_{-\tau}^{\tau} y^2 dF_n(y) \\
&\qquad + \big[1 - (\tau+|\beta_n|)^2 - F_n(\tau) + F_n(-\tau)\big] n\beta_n^2, \qquad (4.28)
\end{aligned}
$$

using both (4.24) and (4.25).

Finally using (4.26) we get

$$
\begin{aligned}
n \int_{-\tau}^{\tau} \frac{y^2}{(1+y^2)} dF_n(y) &\le (1+\tau^2)[1+(\tau+|\beta_n|)^2] n \int_{-\tau}^{\tau} \frac{y^2}{(1+y^2)} dF_n(y) \\
&\qquad + (1+\tau^2)\big[1 - (\tau+|\beta_n|)^2 - F_n(\tau) + F_n(-\tau)\big] n\beta_n^2.
\end{aligned}
\qquad (4.29)
$$

For the type of Lévy process that we consider, the sequence of constants $\{A_n\}$ must necessarily be convergent. Hence $A_n = n\beta_n$ are bounded. This in turn implies that $\beta_n \to 0$ as $n \to \infty$. However we will revisit this again in the next section. For the moment we use the fact that

$$
n\beta_n^2 \to 0 \quad \text{as} \quad n \to \infty. \qquad (4.30)
$$

Now with the selection of τ in such a way so that $\tau, -\tau \in C(G)$, inequalities (4.27) and (4.29) together with Lemma 4.3 imply that:

$$\begin{aligned}\limsup_n \{n \int_{-\tau}^{\tau} \frac{y^2}{1+y^2} dF_n(y)\} &\leq (1+\tau^2)\big[G(\tau) - G(-\tau)\big] \\ &\leq (1+\tau^2)^2 \liminf_n \{n \int_{-\tau}^{\tau} \frac{y^2}{1+y^2} dF_n(y)\}\end{aligned} \tag{4.31}$$

Now applying Lemma 4.4 with $0 < \tau < s$ we get:

$$\begin{aligned}&\lim_{n\to\infty} \left\{n \int_{-\infty}^{s} \frac{y^2}{1+y^2} dF_n(y) - n \int_{-\infty}^{\tau} \frac{y^2}{1+y^2} dF_n(y)\right\} = G(s) - G(\tau) \\ \Rightarrow &\lim_{n\to\infty} \left\{n \int_{-\infty}^{s} \frac{y^2}{1+y^2} dF_n(y) - n \int_{-\tau}^{\tau} \frac{y^2}{1+y^2} dF_n(y)\right\} = G(s) - G(\tau) \\ &\qquad + \lim_{n\to\infty} \left\{n \int_{-\infty}^{-\tau} \frac{y^2}{1+y^2} dF_n(y)\right\} \\ \Rightarrow &\lim_{n\to\infty} \left\{n \int_{-\infty}^{s} \frac{y^2}{1+y^2} dF_n(y)\right\} = G(s) - G(\tau) + \lim_{n\to\infty} \left\{n \int_{-\tau}^{\tau} \frac{y^2}{1+y^2} dF_n(y)\right\} \\ &\qquad + G(-\tau).\end{aligned}$$

Now applying (4.31) we get:

$$\liminf_{n\to\infty} \left\{n \int_{-\infty}^{s} \frac{y^2}{1+y^2} dF_n(y)\right\} \geq G(s) - G(\tau) + \frac{1}{1+\tau^2}\{G(\tau) - G(-\tau)\} + G(-\tau).$$

Taking limit on both sides as $\tau \to 0$ we get:

$$\liminf_{n\to\infty} \left\{n \int_{-\infty}^{s} \frac{y^2}{1+y^2} dF_n(y)\right\} \geq G(s), \qquad s \in C(G).$$

Similarly it follows that:

$$\limsup_{n\to\infty} \left\{n \int_{-\infty}^{s} \frac{y^2}{1+y^2} dF_n(y)\right\} \leq G(s), \qquad s \in C(G).$$

Hence the proof is complete. □

Theorem 4.2 implies that

$$\tilde{G}_n(s) = \sum_{k=1}^{n} \mathbb{E}\Big[\frac{(Y_k^n)^2}{1+(Y_k^n)^2} \mathbb{I}_{[Y_k^n \leq s]}\Big] \longrightarrow G(s), \qquad s \in C(G),$$

as $n \to \infty$. Since the Y_k^n's are identically distributed for each k we have:

$$\tilde{G}_n(s) = n\mathbb{E}\Big[\frac{(Y_k^n)^2}{1+(Y_k^n)^2} \mathbb{I}_{\{Y_k^n \leq s\}}\Big] \longrightarrow G(s), \qquad s \in C(G), \tag{4.32}$$

as $n \to \infty$.

Now by the strong law of large numbers:

$$G^*_{N,n}(s) = \frac{1}{N}\sum_1^{nN} \frac{(Y^n_k)^2}{1+(Y^n_k)^2}\mathbb{I}_{[Y^n_k \le s]} \to \tilde{G}_n(s), \tag{4.33}$$

with probability one, as $N \to \infty$, for every value of s and for every fixed n. Since $G(s)$ and $G^*_{N,n}(s)$ are non-decreasing in s we have:

$$\mathbf{P}\big\{ \lim_{n\to\infty}\lim_{N\to\infty} G^*_{N,n}(s) = G(s) \ \text{ for all } s \in C(G)\big\} = 1.$$

Thus the estimator to be used for $G(s)$ is $G^*_{N,n}(s)$, as defined in (4.33).

4.3.2 On the convergence of β_n

In this section, assume a particular form of the Y^n_k's as:

$$Y^n_k = \frac{A_k - \mu}{\sigma\sqrt{n}}, \qquad \text{where } \mu = \mathbb{E}(A_1) \text{ and } \sigma^2 = \mathbb{V}(A_1).$$

By the central limit theorem we have

$$\sum_{k=1}^{n} Y^n_k \xrightarrow{D} N(0,1), \qquad \text{as } n \to \infty.$$

Now by Tchebychev's inequality, for any $a > 0$,

$$\begin{aligned} \mathbf{P}\big(|A_1-\mu| > \sigma a\sqrt{n}\big) &= \mathbf{P}\big(|A_1-\mu|\mathbb{I}_{\{|A_1-\mu|\ge\sigma a\sqrt{n}\}} > \sigma a\sqrt{n}\big) \\ &\le \frac{\mathbb{E}\big[|A_1-\mu|^2\mathbb{I}_{\{|A_1-\mu|\ge\sigma a\sqrt{n}\}}\big]}{\sigma^2a^2n}. \end{aligned} \tag{4.34}$$

Furthermore,

$$\sigma a\sqrt{1} < \sigma a\sqrt{2} < \cdots \Rightarrow \mathbb{I}_{\{|A_1-\mu|\le\sigma a\sqrt{1}\}} < \mathbb{I}_{\{|A_1-\mu|\le\sigma a\sqrt{2}\}} \cdots \xrightarrow{n\to\infty} \mathbb{I}_{\{|A_1-\mu|\le\infty\}}.$$

Hence by the monotone convergence theorem

$$\mathbb{E}\big[|A_1-\mu|^2\mathbb{I}_{\{|A_1-\mu|<\sigma a\sqrt{n}\}}\big] \to \mathbb{E}\big[|A_1-\mu|^2\big] = \sigma^2.$$

So

$$\begin{aligned} \gamma_n(a) &:= \frac{1}{\sigma^2a^2}\mathbb{E}\big[|A_1-\mu|^2\mathbb{I}_{\{|A_1-\mu|\ge\sigma a\sqrt{n}\}}\big] \\ &= \frac{1}{\sigma^2a^2}\mathbb{E}\big[|A_1-\mu|^2\mathbb{I}_{\{|A_1-\mu|\le\infty\}} - |A_1-\mu|^2\mathbb{I}_{\{|A_1-\mu|<\sigma a\sqrt{n}\}}\big] \\ &= \frac{1}{\sigma^2a^2}\Big[\sigma^2 - \mathbb{E}\big[|A_1-\mu|^2\mathbb{I}_{\{|A_1-\mu|<\sigma a\sqrt{n}\}}\big]\Big], \end{aligned}$$

and hence

$$\gamma_n(a) \longrightarrow 0, \qquad \text{as } n \to \infty.$$

Also (4.34) implies that

$$\mathbf{P}\big(|A_1 - \mu| > \sigma a\sqrt{n}\big) \le \frac{\gamma_n(a)}{n}. \tag{4.35}$$

Now if $M_n = \max\{|Y_1^n|, \cdots, |Y_n^n|\}$ then

$$\begin{aligned} \mathbf{P}\big(M_n \le a\big) &= \mathbf{P}\big(|Y_1^n| \le a, \cdots, |Y_n^n| \le a\big) \\ &= \Big[\mathbf{P}\big(|Y_1^n| \le a\big)\Big]^n \\ &= \Big[\mathbf{P}\big(|A_1 - \mu| \le \sigma a\sqrt{n}\big)\Big]^n \\ &\ge \left[1 - \frac{\gamma_n(a)}{n}\right]^n \longrightarrow e^0 = 1, \qquad \text{as } n \to \infty. \end{aligned}$$

This implies that

$$\mathbf{P}\big(|M_n| > \epsilon\big) = 1 - \mathbf{P}\big(|M_n| \le \epsilon\big) \overset{n\to\infty}{\longrightarrow} 0, \qquad \text{for all } \epsilon > 0.$$

So $M_n \to 0$ in probability. Thus

$$\beta_n := \int_{-r}^{r} y dF_{n,1}(y) \longrightarrow 0, \qquad \text{as } n \to \infty,$$

for $r > 0$.

4.3.3 Relation of G and ν

In this subsection we discuss the connection between G and ν. This relation leads to avail an estimate of ν from that of G.

We can split $\log \Phi_{X_t}$ in (4.1) to write it as:

$$\begin{aligned} \log \Phi_{X_t}(s) &= iats + t\int_{\mathbb{R}\setminus\{0\}} \left[e^{isx} - 1 - \frac{isx}{(1+x^2)}\right] \frac{(1+x^2)}{x^2}\, dG(x) \\ &\quad + t\int_{\{0\}} \left[e^{isx} - 1 - \frac{isx}{(1+x^2)}\right] \frac{(1+x^2)}{x^2}\, dG(x)\,. \end{aligned} \tag{4.36}$$

Now applying l'Hospital's rule, as in (4.9), we see that the integrand

$$\lim_{x\to 0} \left[e^{isx} - 1 - \frac{isx}{(1+x^2)}\right] \frac{(1+x^2)}{x^2} = \frac{-s^2}{2}$$

Hence the second integral in (4.36) reduces to:

$$\int_{\{0\}} \Big[e^{isx} - 1 - \frac{isx}{(1+x^2)}\Big] \frac{(1+x^2)}{x^2} dG(x) = \frac{-s^2}{2}\big[G(0+) - G(0-)\big]. \tag{4.37}$$

Further, we observe that

$$\begin{aligned}
&\Big[e^{isx} - 1 - \frac{isx}{(1+x^2)}\Big] \frac{(1+x^2)}{x^2} \\
= \;& \frac{(e^{isx}-1)(1+x^2)}{x^2} - \frac{is}{x} \\
= \;& \Big(e^{isx} - 1 - isx\mathbb{I}_{[|x|<1]}\Big)\frac{(1+x^2)}{x^2} - \frac{is}{x} + isx\mathbb{I}_{[|x|<1]}\frac{(1+x^2)}{x^2} \\
= \;& \Big(e^{isx} - 1 - isx\mathbb{I}_{[|x|<1]}\Big)\frac{(1+x^2)}{x^2} - \frac{is}{x} + is\mathbb{I}_{[|x|<1]}\frac{1}{x} + isx\mathbb{I}_{[|x|<1]} \\
= \;& \Big(e^{isx} - 1 - isx\mathbb{I}_{[|x|<1]}\Big)\frac{(1+x^2)}{x^2} + \frac{is}{x}\big[\mathbb{I}_{[|x|<1]} - 1\big] + isx\mathbb{I}_{[|x|<1]}.
\end{aligned}$$

Hence the first integral in (4.36) becomes:

$$\begin{aligned}
&\int_{\mathbb{R}\setminus\{0\}} \Big[e^{isx} - 1 - \frac{isx}{(1+x^2)}\Big] \frac{(1+x^2)}{x^2} dG(x) \\
= \;& \int_{\mathbb{R}\setminus\{0\}} \Big(e^{isx} - 1 - isx\mathbb{I}_{[|x|<1]}\Big)\frac{(1+x^2)}{x^2} dG(x) + is\int_{\mathbb{R}\setminus\{0\}} \frac{1}{x}\big[\mathbb{I}_{[|x|<1]} - 1\big]\, dG(x) \\
&\qquad + is\int_{\mathbb{R}\setminus\{0\}} x\mathbb{I}_{[|x|<1]} dG(x) \\
= \;& \int_{\mathbb{R}\setminus\{0\}} \Big(e^{isx} - 1 - isx\mathbb{I}_{[|x|<1]}\Big)\frac{(1+x^2)}{x^2} dG(x) - is\int_{\mathbb{R}\setminus(-1,1)} \frac{1}{x} dG(x) \\
&\qquad + is\int_{(-1,1)\setminus\{0\}} x dG(x).
\end{aligned}$$

Therefore we can rewrite (4.1) as:

$$\begin{aligned}
&\Phi_{X_t}(s) \\
&= \exp\Big\{iats + t\int_{\mathbb{R}\setminus\{0\}} \Big[e^{isx} - 1 - isx\mathbb{I}_{[|x|<1]}\Big]\frac{(1+x^2)}{x^2} dG(x) - ist\int_{\mathbb{R}\setminus(-1,1)} \frac{1}{x} dG(x)\Big\} \\
&\qquad \exp\Big\{ist\int_{(-1,1)\setminus\{0\}} x dG(x) - \frac{s^2}{2}\Big[G(0+) - G(0-)\Big]t\Big\} \\
&= \exp\Big\{is\Big[a - \int_{\mathbb{R}\setminus(-1,1)} \frac{1}{x} dG(x) + \int_{(-1,1)\setminus\{0\}} x dG(x)\Big]t - \frac{s^2}{2}\Big[G(0+) - G(0-)\Big]t\Big\} \\
&\qquad \exp\Big\{t\int_{\mathbb{R}\setminus\{0\}} \Big[e^{isx} - 1 - isx\mathbb{I}_{[|x|<1]}\Big]\frac{(1+x^2)}{x^2} dG(x)\Big\}. \qquad (4.38)
\end{aligned}$$

Comparing (4.38) with (2.4) we see that for the characterization of Lévy–Khinchine's formula in terms of G, as in (4.1), the drift is $\Big[a - \int_{\mathbb{R}\setminus(-1,1)} \frac{1}{x} dG(x) + \int_{(-1,1)\setminus\{0\}} x dG(x)\Big]$, the Gaussian component is $\Big[G(0+) - G(0-)\Big]$ and the explicit relation of G and ν is $\nu(dx) = \frac{(1+x^2)}{x^2}dG(x)$.

Finally in case of subordinator, (4.38) reduces to:

$$\Phi_{X_t}(s) = e^{\left\{is\left[a-\int_{\mathbb{R}\setminus(-1,1)}\frac{1}{x}dG(x)+\int_{(-1,1)\setminus\{0\}}xdG(x)\right]t+t\int_{\mathbb{R}\setminus\{0\}}\left[e^{isx}-1-isx\mathbb{I}_{[|x|<1]}\right]\frac{(1+x^2)}{x^2}dG(x)\right\}}.$$

G being non-decreasing and its estimator being $G^*_{N,n}(s) = \frac{1}{N}\sum_1^{nN}\frac{(Y_k^n)^2}{1+(Y_k^n)^2}\mathbb{I}_{[Y_k^n\le s]}$, as in (4.33), it now follows that an estimator of

$$\nu[\epsilon,\infty) = \int_\epsilon^\infty \nu(dx) = \int_\epsilon^\infty \frac{(1+y^2)}{y^2}dG(y)$$

can be obtained as

$$\begin{aligned}\hat{\nu}[\epsilon,\infty) &= \frac{1}{N}\sum_1^{nN}\frac{1+(Y_k^n)^2}{(Y_k^n)^2}\frac{(Y_k^n)^2}{1+(Y_k^n)^2}\mathbb{I}_{[Y_k^n\le\infty]} - \frac{1}{N}\sum_1^{nN}\frac{1+(Y_k^n)^2}{(Y_k^n)^2}\frac{(Y_k^n)^2}{1+(Y_k^n)^2}\mathbb{I}_{[Y_k^n\le\epsilon]}\\ &= \frac{1}{N}\sum_1^{nN}\mathbb{I}_{[Y_k^n\le\infty]} - \frac{1}{N}\sum_1^{nN}\mathbb{I}_{[Y_k^n\le\epsilon]}. \end{aligned} \qquad (4.39)$$

4.4 Estimating λ_ϵ

In this section we discuss an estimation procedure for the rate λ_ϵ of the Poisson process associated with the jumps larger then a certain, predetermined, threshold $\epsilon > 0$. There may be several ways to get such estimates. We solely present one based on maximizing the likelihood function.

Assume that we observe the process at equally spaced time points $\{t_i\}_{i=1}^n$ such that $t_i - t_{i-1}$ is constant. Further assume that $n_1 \le n_2 \le \cdots$ is a set of non-negative integers, including zero, such that the n_i are the number of jumps larger than the threshold $\epsilon > 0$, at some observation point, not necessarily at the i^{th} observation point. Then a typical set of observations may be as follows: nodes $t_1^{n_1}, t_2^{n_1}, \cdots, t_{j_{n_1}}^{n_1}$ corresponds to n_1 jumps larger than threshold, nodes $t_1^{n_2}, t_2^{n_2}, \cdots, t_{j_{n_2}}^{n_2}$ corresponds to n_2 jumps larger than threshold and so on. Where

$$\sum_i \#\{t^{n_i}\ ;\ \text{nodes (time points) which accounts for } n_i \text{ jumps larger than } \epsilon\} = n.$$

Then the corresponding likelihood function can be written as:

$$L(\lambda_\epsilon) = \left(\frac{e^{-\lambda_\epsilon}\lambda_\epsilon^{n_1}}{n_1!}\right)^{j_{n_1}} \left(\frac{e^{-\lambda_\epsilon}\lambda_\epsilon^{n_2}}{n_2!}\right)^{j_{n_2}} \cdots \left(\frac{e^{-\lambda_\epsilon}\lambda_\epsilon^{n_i}}{n_i!}\right)^{j_{n_i}} \tag{4.40}$$

where n_i's and j_{n_i}'s are need to be computed from data.

4.5 Claim Size Distribution Through $\hat{\nu}$

In this section we discuss the distribution of jumps which are larger than a certain threshold ϵ. We will see how we can approximate such a distribution using our estimator of ν.

The conditional distribution of jump-sizes ΔX given that $\Delta X > \epsilon$, can be obtained as:

$$\begin{aligned} F^{\epsilon}_{\Delta X}(x) &= \mathbf{P}(\Delta X \leq x \mid \Delta X > \epsilon) \\ &= \frac{\int_\epsilon^x f(y)dy}{1 - F(\epsilon)}, \qquad \text{where } F' = f \\ &= \frac{F(x) - F(\epsilon)}{\bar{F}(\epsilon)}, \qquad \text{and } \bar{F} = 1 - F. \end{aligned} \tag{4.41}$$

Now from the relation between ν and F we have that $\nu(0,x) = \lambda F(x)$, where $\lambda = \int_0^\infty \nu(dx) = \nu(0,\infty)$. So

$$\begin{aligned} \bar{F}(x) &= 1 - \frac{\nu(0,x)}{\lambda} \\ &= \frac{\lambda - \{\nu(0,\infty) - \nu(x,\infty)\}}{\lambda} \\ &= \frac{\nu(x,\infty)}{\lambda}. \end{aligned} \tag{4.42}$$

Thus from (4.41),

$$F^{\epsilon}_{\Delta X}(x) = \frac{\lambda\{F(x) - F(\epsilon)\}}{\nu(\epsilon,\infty)} \tag{4.43}$$

However using the relation between ν and F we have:

$$\begin{aligned} F(x) - F(\epsilon) &= \frac{1}{\lambda}\{\nu(0,x) - \nu(0,\epsilon)\} \\ &= \frac{1}{\lambda}\Big[\nu(0,\infty) - \nu(0,x) - \{\nu(0,\infty) - \nu(0,\epsilon)\}\Big] \\ &= \frac{1}{\lambda}\{\nu(\epsilon,\infty) - \nu(x,\infty)\}. \end{aligned} \tag{4.44}$$

Then (4.43) turns into:

$$F^{\epsilon}_{\Delta X}(x) = \frac{\nu(\epsilon,\infty) - \nu(x,\infty)}{\nu(\epsilon,\infty)}. \tag{4.45}$$

Clearly from (4.43),

$$F^{\epsilon}_{\Delta X}(\epsilon) = \frac{\lambda\{F(\epsilon) - F(\epsilon)\}}{\nu(\epsilon, \infty)} = 0$$

and

$$F^{\epsilon}_{\Delta X}(\infty) = \frac{\lambda\{F(\infty) - F(\epsilon)\}}{\nu(\epsilon, \infty)} = 1, \qquad \text{considering (4.42).}$$

So $F^{\epsilon}_{\Delta X}(x)$ is a distribution on (ϵ, ∞).

In practice we do not know from which distribution our claims arise. We just know a finite set of values of claims which are larger then some threshold ϵ. Equation (4.45) then helps us to obtain a purely non-parametric approximation of the underlying distribution of such claims through the estimator of ν.

Chapter 5

Numerical Implementation

In this chapter we test our estimators numerically for estimating the intensity rate λ_ϵ of jumps exceeding the threshold ϵ and for approximating the conditional jump size distribution of jumps larger than ϵ. Finally we compare our approximation with the theoretical and estimated distributions of such jumps.

Observations are sampled from simulated aggregate claim subordinator paths. We simulate gamma, inverse Gaussian and α-stable subordinators on a time interval of $[0, 500]$.

We sample the path at nodes spaced equally by 50 time units, that is we sample at time points $50i$, for $i = 1 \cdots 10$. Then we estimate the average number of jumps, λ_ϵ, larger than a given threshold ϵ, occurring between two consecutive nodes. We use equation (4.40) for this estimation.

To get an estimate of the Lévy measure $\nu(\epsilon, \infty)$ of a subordinator $X = \{X_t\, ;\, t \geq 0\}$ with jump exceeding the threshold , we use (2.24) and (2.25). Note that, by Theorem 3.1, to such a Lévy process there corresponds a Poisson point process of rate λ. Now recall our estimate of rate λ_ϵ is a per node estimate, where the nodes are at time points $50i$. Instead if our program uses only one node at time point 500, our rate estimate is just the number of jumps in $[0, 500]$ which exceed the threshold ϵ. This rate is the same Poisson rate appearing in (2.24). Thus we obtain:

$$|500 - 0| \int_\epsilon^\infty g(x)dx = \sharp\big\{t\, ;\, 0 \leq t \leq 500,\ \triangle X_t \in (\epsilon, \infty)\big\}.$$

The number on the right hand side is an output of our program for any simulated path of a subordinator. Thus we have

$$\hat{\nu}(\epsilon,\infty)=\frac{\sharp\{t\mid 0\le t\le 500,\ \triangle X_t\in(\epsilon,\infty)\}}{500}. \tag{5.1}$$

We then apply (4.39) to estimate the Lévy measure through the estimator of G. We compare these two estimators of the Lévy measure.

These estimates are consistent with the definition of Lévy measure of a Lévy process appearing in Definition 2.1.

We produce four plots for each model, with their estimated λ_ϵ and ν, corresponding to different thresholds. The decreasing threshold sizes show the convergence of the truncated jump sample paths to the original non-truncated jump paths which agrees with intuition.

5.1 Estimation of the Rate λ_ϵ

In this section we use our estimators of Lévy measure to estimate the rate λ_ϵ of the aggregate claims process. We produce such estimates for both gamma and inverse Gaussian subordinators.

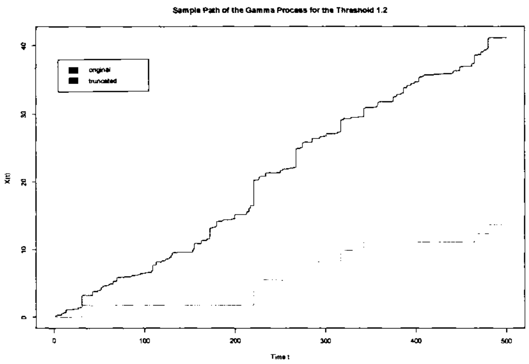
5.1.1 Gamma process

The chosen simulation parameters for the gamma process are $\alpha = 0.09/20$ and $\beta = 1$. These parameter values make the sample paths easily observable at different jump sizes. However one can rescale using the fact that for a gamma(1,1) process $X = \{X_t\,;\,t\ge 0\}$, the process $X^* = \{\alpha X_{\beta t}\,;\,t\ge 0\}$ is gamma(α,β). A useful reference for simulating gamma and variance gamma processes is [2].

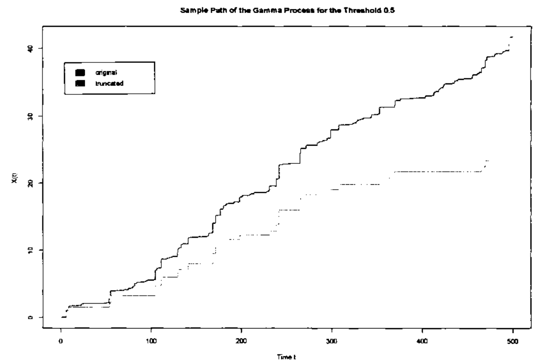
We produce four sample paths of the gamma aggregate claims process using the same parameters above. The paths of threshold exceeding aggregate claims process corresponding to different thresholds appears with each path. All the graphs appear in Figure 5.1. The corresponding estimates are summarized in Table 5.1.

5.1.2 Inverse Gaussian process

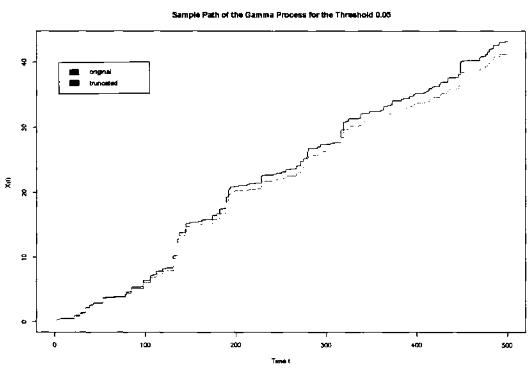
In this section we apply our estimators on inverse Gaussian aggregate claims process for threshold exceeding claims. The chosen simulation parameters are $\mu = 200$ and $\theta = 0.3$.



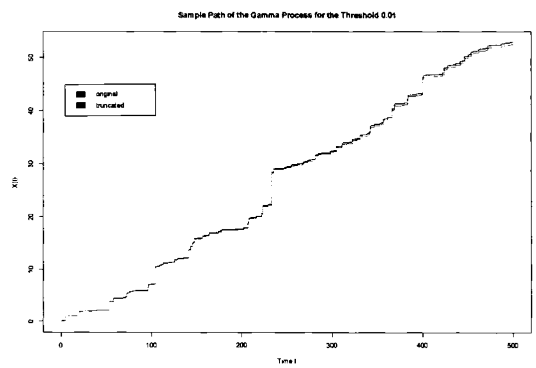
(a) Threshold 1.2



(b) Threshold 0.5



(c) Threshold 0.05



(d) Threshold 0.01

Figure 5.1: Gamma ($\alpha = \frac{0.09}{20}$, $\beta = 1$) and truncated gamma processes for the thresholds $\epsilon = 1.2, 0.5, 0.05$ and 0.01.

However we mention that if $X \sim IG(\mu, \theta)$ then, for any $c > 0$, $cX \sim IG(\sqrt{c}\mu, \frac{\theta}{\sqrt{c}})$. This rescaling property can be used to model different jump sizes. Figure 5.2 reports four simulated paths for the inverse Gaussian aggregate claims processes together with the corresponding threshold exceeding paths. Estimates are summarized in Table 5.2.

5.1.3 α-stable process

We apply our estimators on α-stable aggregate claims process for different thresholds. Here we mention that simulating positive α-stable process is a bit involved. A very useful reference is [18]. We simulate standard α-stable process for $\alpha = 0.8$ and $\beta = 1$. However we

Table 5.1: Estimates of rate λ_ϵ for gamma ($\alpha = \frac{0.09}{20}$, $\beta = 1$) process.

Threshold (ϵ)	$\hat{\lambda}_\epsilon$ based on (5.1)	$\hat{\lambda}_\epsilon$ based on (4.39)	% error
1.2	0.7	0.625	10.71
0.5	2.5	2.625	5.00
0.05	10.8	10.625	1.60
0.01	18.2	18.00	1.11

Table 5.2: Estimates of rate λ_ϵ for IG ($\mu = 200$, $\theta = 0.3$) process.

Threshold (ϵ)	$\hat{\lambda}_\epsilon$ based on (5.1)	$\hat{\lambda}_\epsilon$ based on (4.39)	% error
40000	0.7	0.6	14.280
20000	2.7	2.75	3.703
5000	4.5	4.35	3.340
100	43.3	43.0	0.690

mention that for a standard α-stable random variable $X \sim S_\alpha(1, \beta, 0)$, the random variable

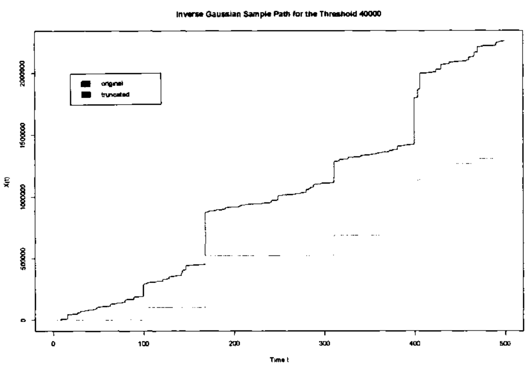
$$Y = \begin{cases} \sigma X + \mu & \alpha \neq 1 \\ \sigma X + \frac{2}{\pi}\beta\sigma \log \sigma + \mu & \alpha = 1 \end{cases}$$

is $S_\alpha(\sigma, \beta, \mu)$, where σ, μ and β are respectively scaling , location and skewness parameters and α is usually known as index which appears in the characteristic exponent. This rescaling property can be used for simulating general α-stable process. For general properties and different characterizations of stable distributions we refer to [21] and [20].

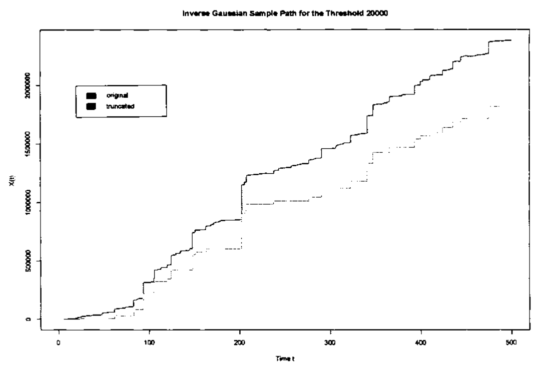
In Figure 5.3 we produce four simulated paths together with their threshold exceeding paths for different thresholds. Corresponding estimates are summarized in Table 5.3.

5.2 Approximation of Jump Size Distribution

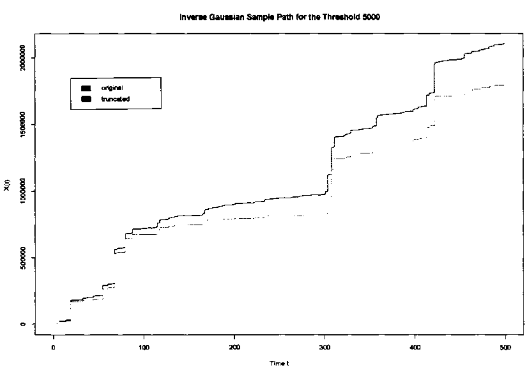
Here we approximate numerically the jump size distribution of jumps larger than some threshold ϵ. In this regard we use the theoretical developments in Chapter 4. Graphical approximations of cumulative distribution functions (CDF) and probability density functions (PDF) of such jumps, corresponding to different thresholds, are produced. Finally a



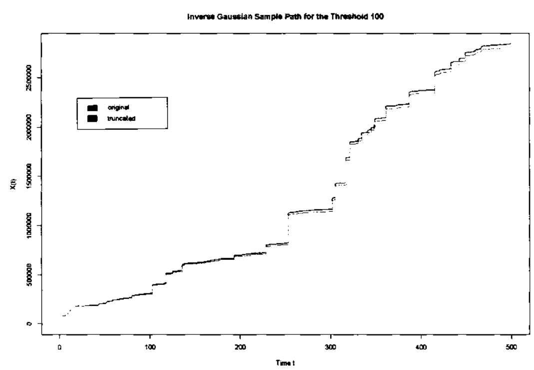
(a) Threshold 40000



(b) Threshold 20000



(c) Threshold 5000



(d) Threshold 100

Figure 5.2: Inverse Gaussian ($\mu = 200$, $\theta = 0.3$) and truncated processes for thresholds $\epsilon = 40000, 20000, 5000$ and 100.

goodness of fit test is conducted.

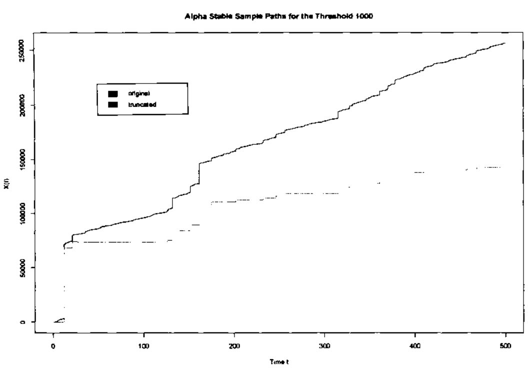
5.2.1 Gamma process

At first we produce different estimates of the parameters which will be extensively used in the sequel. Table 5.4 summarizes these estimates.

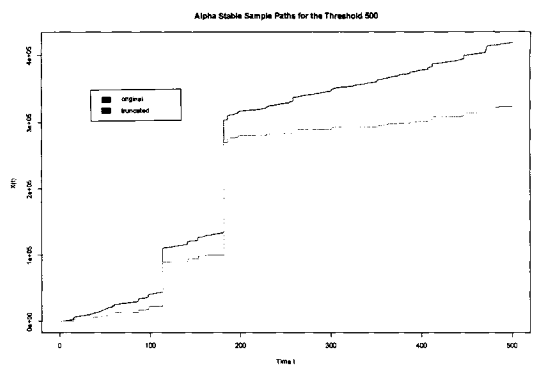
The corresponding densities for each set of estimates appear in Figure 5.4.

Figure 5.5 shows the corresponding CDF's together with our non-parametric approximation of the empirical CDF through the estimator of Lévy measure ν.

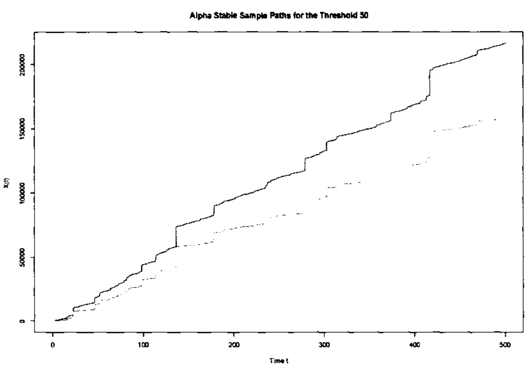
Finally in Table 5.5, we show the goodness of fit of such approximations.



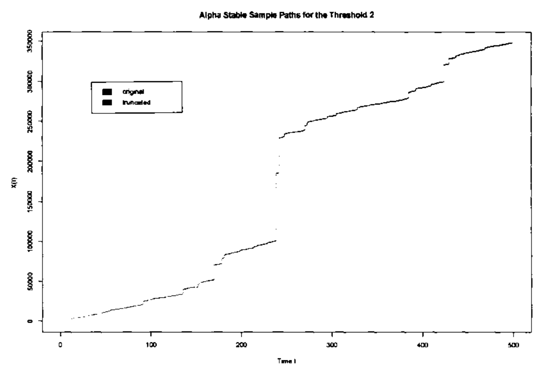
(a) Threshold 1000



(b) Threshold 500



(c) Threshold 50



(d) Threshold 2

Figure 5.3: α-stable and truncated processes for thresholds $\epsilon = 1000, 500, 50, 2$.

5.2.2 Inverse Gaussian process

We reproduce the estimates of parameters in Table 5.6. These parameter estimates are used in the sequel.

The corresponding densities for each set of estimates appear in Figure 5.6.

Figure 5.7 shows the corresponding CDF's together with our non-parametric approximation of the empirical CDF using the estimator of Lévy measure ν.

Table 5.7 shows the goodness of fit.

We conclude by producing a comparative graphical study. It shows that when the claims arise from a gamma distribution, conditioned by the fact that they are larger than some threshold ϵ, then the inverse Gaussian distribution, conditioned on the same fact that claims are larger than the same threshold, provides a poor fit to the true distribution. Figure 5.8 shows the graphs supporting this claim.

Table 5.3: Estimates of rate λ_ϵ for α-stable process with $\alpha = 0.8$.

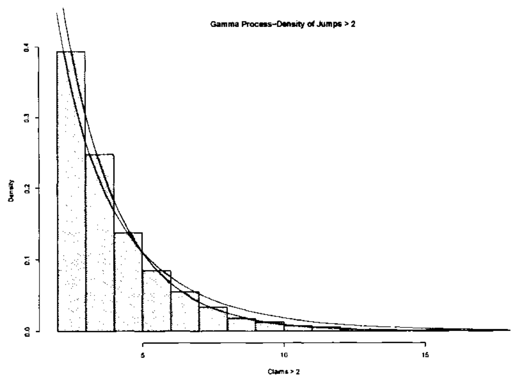
Threshold (ϵ)	$\hat{\lambda}_\epsilon$ based on (5.1)	$\hat{\lambda}_\epsilon$ based on (4.39)	% error
1000	2.1	2.25	7.142857
500	5.0	5.25	5.000000
50	34.8	35.0	0.574712
2	953.2	953.125	0.007868

Table 5.4: Estimates of gamma parameters

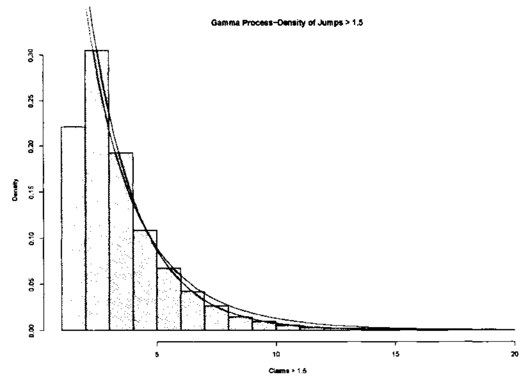
Different estimates			
Threshold	Theoretical (α, β)	MLE $(\hat{\alpha}, \hat{\beta})$	MME $(\tilde{\alpha}, \tilde{\beta})$
2	(1, 0.5)	(0.9169898, 0.4869544)	(0.3752765, 0.2759984)
1.5	(1, 0.5)	(0.9655107, 0.492039)	(0.5301264, 0.3282412)
1	(1, 0.5)	(0.9890272, 0.4900115)	(0.720914, 0.3919935)
0.5	(1, 0.5)	(0.9356638, 0.4733607)	(0.8452734, 0.4399136)

Table 5.5: Goodness of fit test for different CDF's.

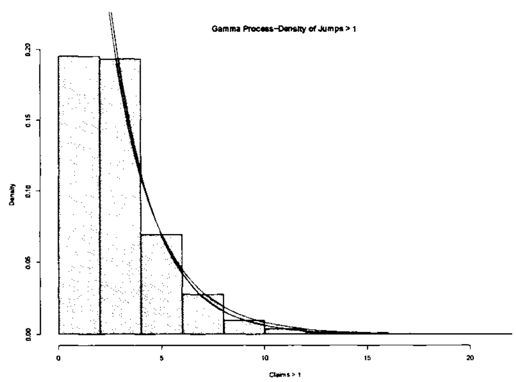
Kolmogorov-Smirnov test				
Threshold	MLE	MME	Non-parametric	KS acceptance limit
2	0.007902474	0.08722181	0.006913213	0.02025472
1.5	0.009275763	0.05835448	0.006510432	0.01778040
1	0.009463755	0.02490687	0.006077613	0.01578701
0.5	0.005906946	0.01011186	0.005315381	0.01385737



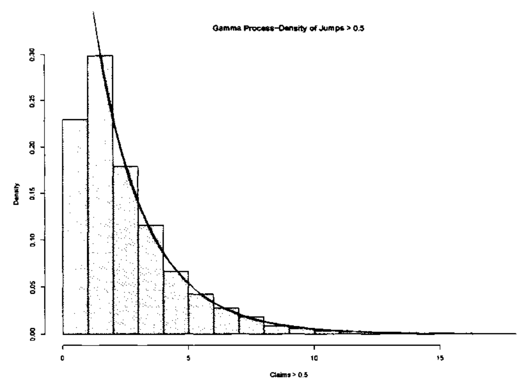
(a) Threshold 2



(b) Threshold 1.5



(c) Threshold 1

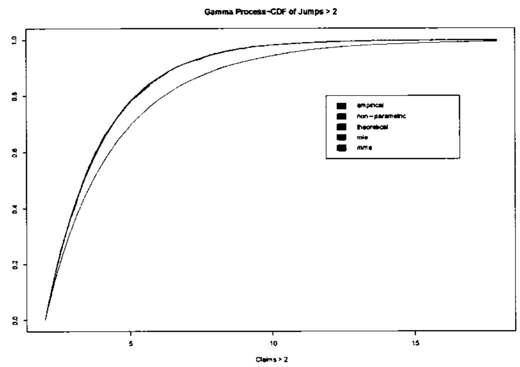


(d) Threshold 0.5

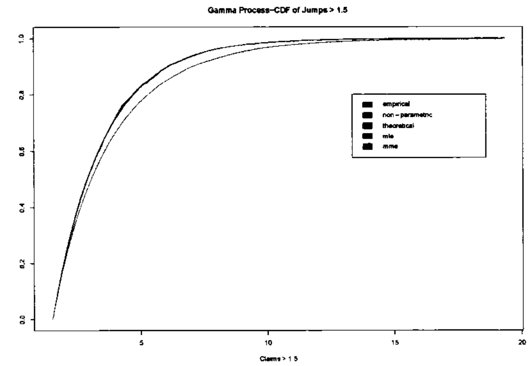
Figure 5.4: Gamma ($\alpha = 1$, $\beta = 0.5$) and truncated processes for thresholds $\epsilon = 2, 1.5, 1$ and 0.5.

Table 5.6: Estimates of inverse Gaussian parameters

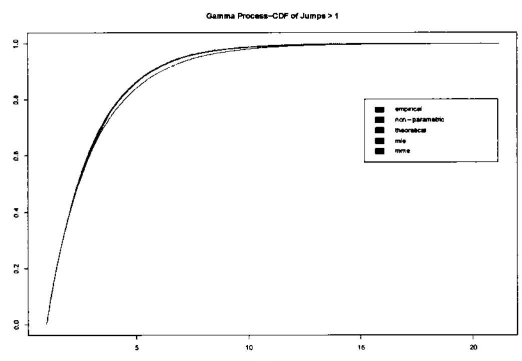
Different estimates			
Threshold	Theoretical (μ, θ)	MLE $(\hat{\mu}, \hat{\theta})$	MME $(\tilde{\mu}, \tilde{\theta})$
250	(40, 5)	(509.7107, 1615.089)	(440.8967, 683.3879)
200	(40, 5)	(456.3663, 1454.848)	(412.7533, 703.2026)
150	(40, 5)	(371.5199, 927.676)	(332.6884, 403.3276)
100	(40, 5)	(278.7732, 580.8884)	(252.227, 293.9069)



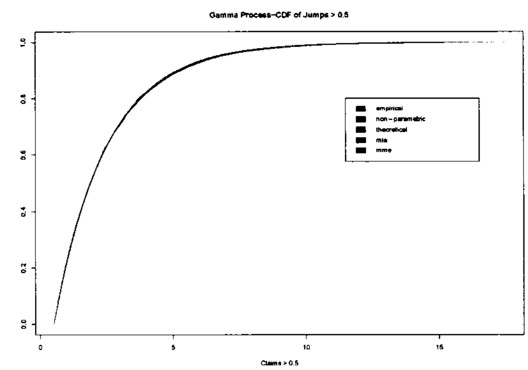
(a) Threshold 2



(b) Threshold 1.5



(c) Threshold 1

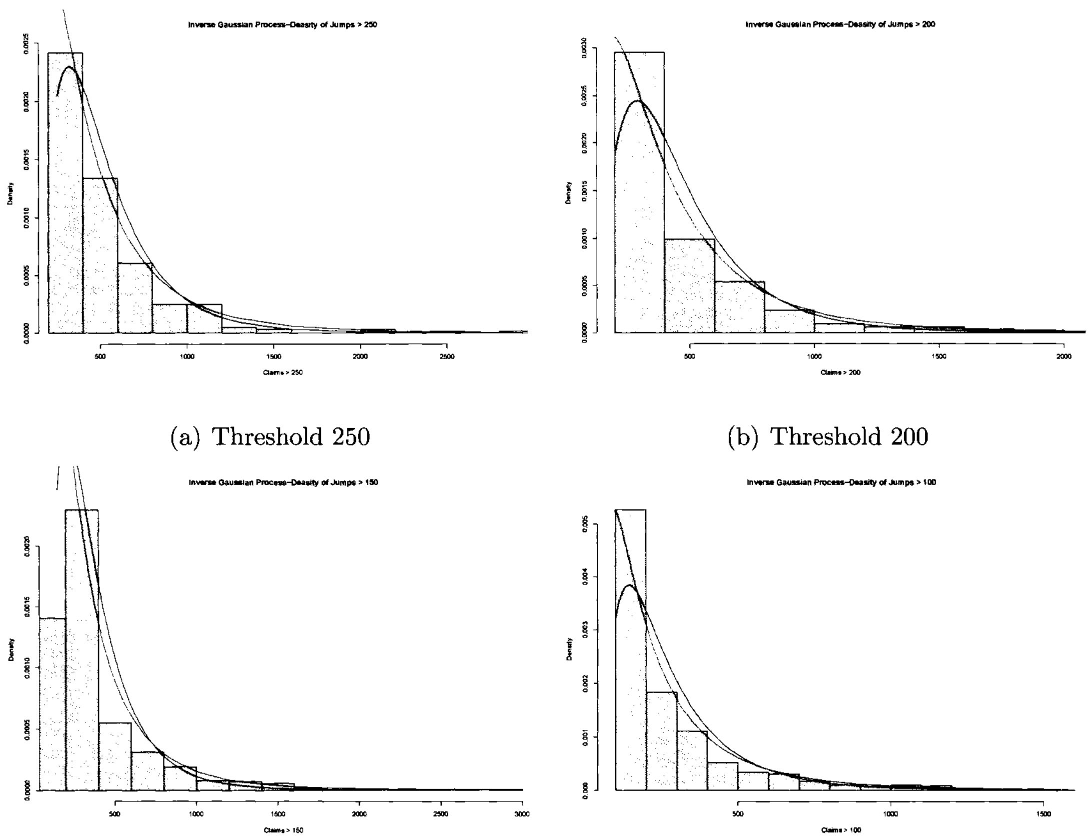


(d) Threshold 0.5

Figure 5.5: Gamma ($\alpha = 1$, $\beta = 0.5$) and truncated processes for thresholds $\epsilon = 2, 1.5, 1$ and 0.5.

Table 5.7: Goodness of fit test for different CDF's.

Kolmogorov–Smirnov test				
Threshold	MLE	MME	Non-parametric	KS acceptance limit
250	0.170295	0.1261079	0.03170745	0.06974276
200	0.1619935	0.1091162	0.02492963	0.05917869
150	0.1792451	0.1357120	0.01654300	0.04883909
100	0.1734932	0.1182062	0.01474935	0.04024409



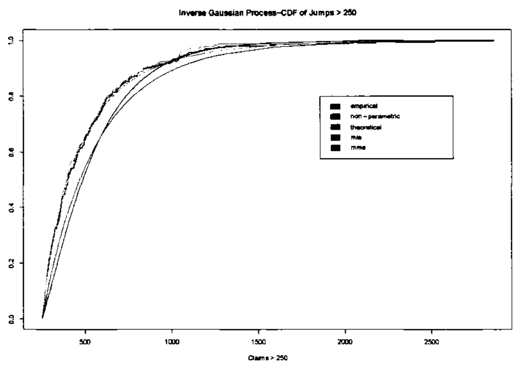
(a) Threshold 250

(b) Threshold 200

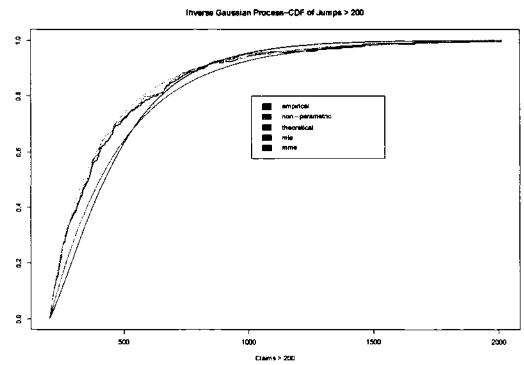
(c) Threshold 150

(d) Threshold 100

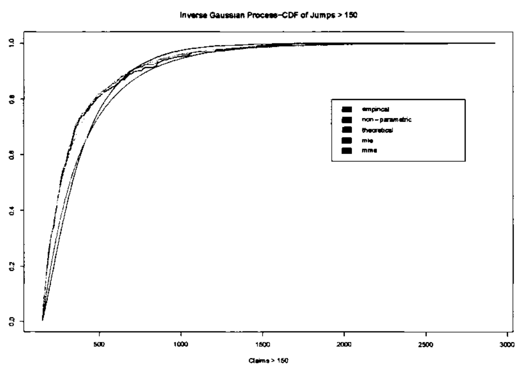
Figure 5.6: Inverse Gaussian (μ = 40, θ = 5) and truncated processes for thresholds $\epsilon = 250, 200, 150$ and 100.



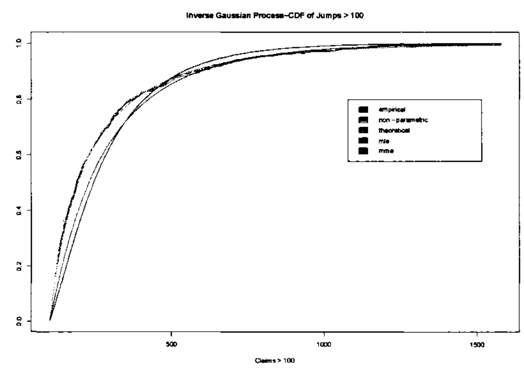
(a) Threshold 250



(b) Threshold 200

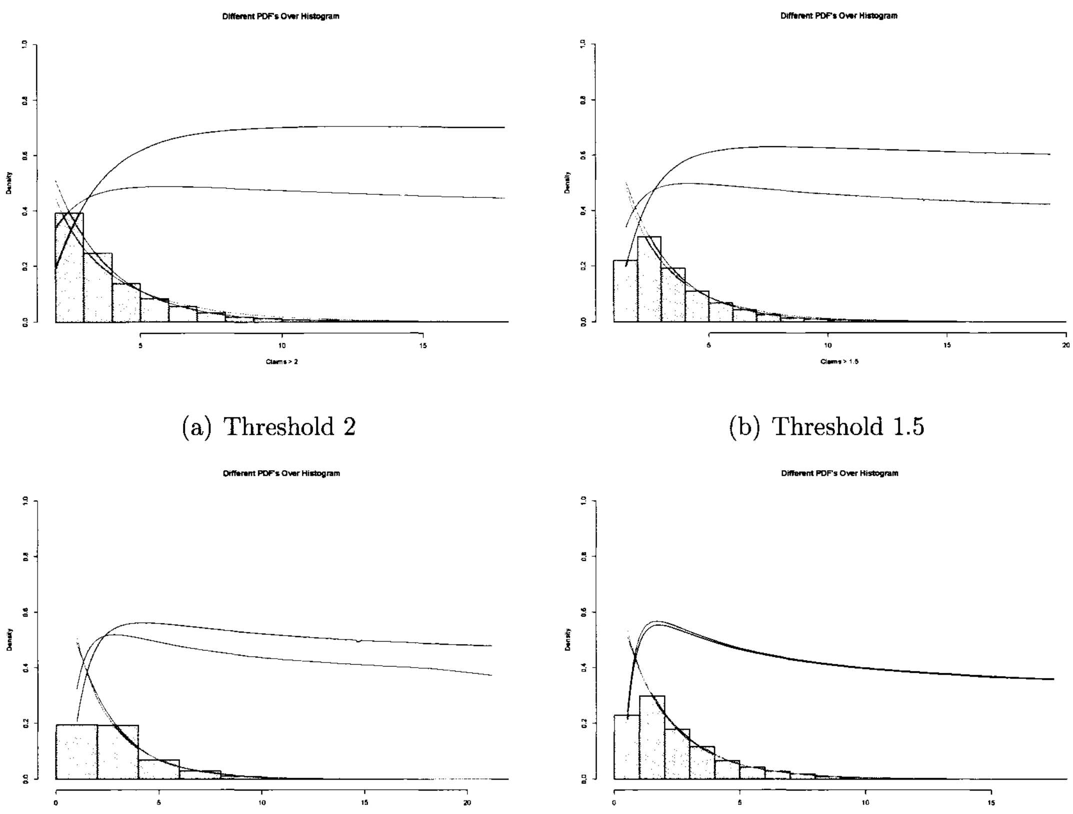


(c) Threshold 150



(d) Threshold 100

Figure 5.7: Inverse Gaussian ($\mu = 40$, $\theta = 5$) and truncated processes for thresholds $\epsilon = 250, 200, 150$ and 100.



(a) Threshold 2 (b) Threshold 1.5

(c) Threshold 1 (d) Threshold 0.5

Figure 5.8: Histogram of ϵ-exceeding gamma claims with estimated conditional gamma and conditional IG densities.

Conclusion

Lévy processes are gaining popularity in actuarial and financial modeling. In this thesis we see how the Lévy measure plays a key role in applications of Lévy processes. In particular, the estimation of the Lévy measure from data is required if Lévy processes are to be used to model claims in Risk Theory.

The relation between classical risk processes and subordinators (increasing Lévy processes) is explored. We obtain results for the jump function G appearing in Lévy's characterization of the Lévy–Khinchine formula. A non–parametric estimator of G is discussed. A detailed relation between G and ν, the Lévy measure, is derived, yielding an estimator of ν. The latter gives an estimator of the Poisson intensity rate λ_ϵ and the claim size distribution F^ϵ for claims larger than the threshold ϵ. Extensive numerical simulations illustrate the paths of gamma, inverse Gaussian and α–stable claim subordinators and their corresponding estimates for λ_ϵ and F^ϵ.

The main results obtained in the thesis and conclusions drawn from them are summarized as follows:

- The relation between the jump measure J_X and the Lévy measure ν is given explicitly, leading to a nice proof.
- The change from Paul Lévy's jump function G to the now commonly used jump measure ν is independent from the choice of cut–off function. It only affects the expressions for the drift parameter a and the Gaussian component b.
- The proposed non–parametric estimator of the Lévy measure is shown to be strongly consistent.
- This non–parametric estimator provides a better fit than parametric ones (MLE or MME) for the true distribution of jumps larger than the threshold.

- The goodness–of–fit results are consistent for different simulated subordinators or parameter values.

- Small threshold values yield more accurate estimations, for both λ_ϵ and F^ϵ. This is consistent with intuition, as smaller threshold values lead to less truncation of the sample path, hence to an estimation based on more observed jumps. In applications, a sufficiently small threshold can be selected to part jumps viewed as insurance claims from the smaller, more numerous jumps, interpreted as other expenses (commissions, taxes, etc.).

Bibliography

[1] Applebaum, David (2004). *Lévy Processes and Stochastic Calculus.* Cambridge University Press: Cambridge, UK.

[2] Avramidis, A. N., L'Ecuyer, P. and Tremblay, Pierre-Alexandre (2003). Efficient simulation of gamma and variance gamma processes. *Proceedings of the 2003 Winter Simulation Conference.*

[3] Bertoin, Jean (1996). *Lévy Processes.* Cambridge University Press: Cambridge, UK.

[4] Bulm, J.R. and Rosenblat, M. (1959). On the structure of infinitely divisible distributions. *Pacific J. Math.* 9(1), 1–7.

[5] Cont, Rama and Tankov, Peter (2004). *Financial Modelling With Jump Processes.* Chapman & Hall/CRC Financial Mathematics Series: London .

[6] Erickson, K. Bruce and Maller, Ross. A. (2007). Finiteness of integrals of functions of Lévy processes. *Proc. London Math. Soc.* 3(94), 386–420.

[7] Feller, W. (1971). *An Introduction to Probability Theory and its Application.* Vol. 2, 2nd ed., Wiley: New York.

[8] Garrido, J. and Morales, M. (2006). On the expected discounted penalty function for Lévy risk processes. *The North American Actuarial Journal,* 10(4), 196–216.

[9] Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley: Cambridge, Mass., USA.

[10] Gut, Allun (2005). *Probability : A Graduate Course.* Springer Texts in Statistics: New York.

[11] Jang, J. (2007). Jump diffusion processes and their applications in insurance and finance. *Insurance: Mathematics and Economics*, 41, 62–70.

[12] Jongbloed, G. and Van Der Meulen, F.H. (2006). Parameter estimation for subordinators and induced OU process. *Scandinavian Journal of Statistics*, 33, 825–847.

[13] Major, Péter. Limit theorems and infinitely divisible distributions. *The mathematical Institute of Hungarian Academy of Sciences*. (Available at http://www.renyi.hu/ major/probability/divisible.html)

[14] Roussas, George, G. (2005). *An Introduction to Measure-Theoretical Probability*. Elsevier Academic Press: Boston.

[15] Rubin, Herman and Tucker, Howard G. (1959). Estimating the parameters of a differential process. *Annals of Mathematical Statistics* 30(3), 641–658.

[16] Sato, Ken–Iti (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press: Cambridge, UK.

[17] Schoutens, W. (2003). *Lévy Processes in Finance*. Wiley: New York.

[18] Weron, R. (1996). On the Chamblers–Mallows–Stuck method for simulating skewed stable random variables. *Statistics and Probability Letters*, 28, 165–171.

[19] Winkel, Matthias (2007). *Lévy Processes and Finance*. Lecture notes, Oxford University, UK. (Available at http://www.stats.ox.ac.uk/ winkel/ms3b.html)

[20] Zolotarev, V. M. (1966). On representation of stable laws by integrals. *Selected Translations in Mathematical Statistics and Probability*, 6, 84–88.

[21] Zolotarev, V. M. (1986). *One–Dimensional Stable Distributions*. American Mathematical Society: Providence, RI, USA.

Appendix A

Preliminaries in Probability

The goal of this appendix is to introduce the basic terminology and first results, some with proof, from the general theory of Lévy processes. For details the reader is referred to [14], [16].

A.1 Topics from Probability Theory

Definition A.1 *Consider a measurable space $(\Omega, \mathcal{F}, \mu)$ where the measure μ is a set function defined as $\mu : \mathcal{F} \longrightarrow \mathbb{R}^+$ satisfying:*

[i] $\forall\, A \in \mathcal{F},\ \mu(A) \geq 0$,

[ii] for any collection of disjoint events $A_j \in \mathcal{F}$, $j = 1, 2, 3 \cdots$

$$\mu\Big[\bigcup_{i=1}^{n} A_j\Big] = \sum_{i=1}^{n} \mu(A_j)\,,$$

[iii] $\mu(\emptyset) = 0$.

μ is called σ—finite if $\mu(\Omega) = \infty$ but there exists a partition of Ω, say $\{A_j\}_{j=1}^{\infty}$ such that $\mu(A_j) < \infty$ for $j = 1, 2, \ldots$. In particular if $\mu(\Omega) = 1$ then μ is called a probability measure and is denoted by $\mathbf{P}$. $(\Omega, \mathcal{F}, \mathbf{P})$ is then called a probability space.

A property, say "S", about the elements of Ω is said to hold almost everywhere, abbreviated as a.e., if the set on which the property does not hold has *measure zero*, i.e.

$$\aleph = \{\omega \in \Omega \mid S(\omega)\,\text{is false}\} \quad \Rightarrow \quad \mu(\aleph) = 0\,.$$

In the case when μ is a probability measure the terminology is rephrased as "almost surely" and is abbreviated as a.s..

Definition A.2 *(Indicator function) Let $\mathcal{N}$ be a set. The indicator function of a subset A of $\mathcal{N}$ is defined as*

$$\mathbb{I}_A : \mathcal{N} \to \{0, 1\},$$

where

$$\mathbb{I}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} . \tag{A.1}$$

A.1.1 Notions of convergence

Let $\{X_n \,;\, n \in \mathbb{N}\}$ be a sequence of $\mathbb{R}$–valued random variables (henceforth r.v.). We say that :

- X_n converges to X *a.s.* if

  $$\lim_{n\to\infty} X_n(\omega) = X(\omega), \quad \forall\, \omega \in \Omega \setminus \aleph,$$

  where $\aleph \in \mathcal{F}$ satisfies $\mathbf{P}(\aleph) = 0$.

- X_n converges to X *in probability* if

  $$\lim_{n\to\infty} \mathbf{P}\big(|X_n(\omega) - X(\omega)| > a\big) = 0, \qquad \forall\, a > o.$$

  In the case when the associated measure is not necessarily a probability measure the convergence is know as *convergence in measure*:

  $$\lim_{n\to\infty} \mu\big(|X_n(\omega) - X(\omega)| > a\big) = 0.$$

- X_n converges to X *in distribution* if

  $$\lim_{n\to\infty} F_n(x) \to F(x) \;\text{ where }\; F_X(x) = \mathbf{P} \circ X^{-1}(-\infty, x]$$

  or equivalently

  $$\lim_{n\to\infty} \int_{\mathbb{R}} f(x) F_{X_n}(dx) \to \int_{\mathbb{R}} f(x) F_X(dx)$$

$\forall f \in \mathbb{C}_b(\mathbb{R})$-the class of all bounded continuous functions in $\mathbb{R}$. This notion of convergence is also known as *weak convergence* of $\{F_n\}$ to F and is important for what follows in the sequel.

Proposition A.1 *If X_n converges to X in probability, $X_n \xrightarrow{p} X$, then the distribution of X_n converges to the distribution of X.*

Definition A.3 (Characteristic function) *Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbf{P})$ taking values in $\mathbb{R}$. Its Characteristic function $\Phi : \mathbb{R} \to \mathbb{C}$is defined by*

$$\begin{aligned} \Phi_X(t) = E(e^{itX}) &= \int_\Omega e^{itX(\omega)} \mathbf{P}(d\omega) \quad \textit{for allt} \in \mathbb{R} \\ &= \int_{\mathbb{R}} e^{itx} dF \\ &= \int_{-\infty}^{\infty} e^{itx} f(x) dx. \end{aligned} \tag{A.2}$$

where $F_X(x) = \mathbf{P}oX^{-1}(-\infty, x]$.

Clearly

$$|\Phi_X(t)| = \left| \int_{-\infty}^{\infty} e^{itx} f(x) dx \right| \leq \int_{-\infty}^{\infty} f(x) dx = 1.$$

Definition A.4 (Fourier transform) *Let f be a complex valued integrable function on $\mathbb{R}$. The Fourier transform $\mathcal{F}(f)$ of f is a complex valued function defined by*

$$\mathcal{F}(f) = \hat{f}(s) = \int_{-\infty}^{\infty} e^{-isx} f(x)\, dx \qquad s \in \mathbb{R}. \tag{A.3}$$

Remark A.1 From (A.2) and (A.3) we see that $\hat{f}(s) = \Phi(-s)$.

Example A.1 *Consider the Gaussian density function*

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}}, \qquad x \in \mathbb{R}.$$

Using (A.2) *its easy to see that the characteristic function of f_X is*

$$\Phi_X(t) = e^{-\frac{1}{2}t^2\sigma^2}, \qquad t \in \mathbb{R}.$$

By the remark above its Fourier transform is given by

$$\hat{f}(s) = e^{-\frac{1}{2}(-s)^2\sigma^2} = e^{-\frac{1}{2}s^2\sigma^2}, \qquad s \in \mathbb{R}.$$

Now for the transformed density corresponding to the centered r.v. $X - \eta$

$$\overline{f}_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\eta)^2}{2\sigma^2}}$$

using the shifting property of Fourier transform we get

$$\hat{\overline{f}}(s) = e^{-is\eta} \cdot \hat{f}(s) = e^{-is\eta} \cdot e^{-\frac{1}{2}s^2\sigma^2}, \qquad s \in \mathbb{R},$$

where $\eta = \mathbb{E}(X)$. *Now using Remark A.1 we get the characteristic function of the transformed density* $\overline{f}_X(x)$ *is given by:*

$$\overline{\Phi}_X(t) = e^{it\eta} \cdot e^{-\frac{1}{2}t^2\sigma^2} = e^{it\eta - \frac{1}{2}t^2\sigma^2} \qquad t \in \mathbb{R}\,.$$

This form of the characteristic function of a Gaussian density occurs frequently in the thesis.

Analogous to the above definition, the joint characteristic function $\Phi_{XY}(t_1, t_2)$ of random variables X and Y is defined as

$$\Phi_{XY}(t_1, t_2) = \mathbb{E}\left[e^{i(t_1X + t_2Y)}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(t_1x+t_2y)} f_{XY}(x, y)dxdy.$$

This can be recognized as a two–dimensional *Fourier transform* of f_{XY} with the appropriate change of sign in the exponent as suggested by Remark A.1. Generalizing this notion the joint characteristic function of n random variables $X_1, X_2, \cdots, X_n$ can be defined as

$$\begin{aligned}\Phi_{X_1,X_2\cdots,X_n}(t_1, t_2, \cdots, t_n) &= \mathbb{E}\left[e^{i(t_1X_1+t_2X_2+\cdots+t_nX_n)}\right] \\ &= \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-times} e^{i(t_1x_1+\cdots+t_nx_n)} f_{X_1,\cdots,X_n}(x_1, \cdots, x_n)dx_1\cdots dx_n.\end{aligned}$$

Using a vector notation

$$\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \quad \text{and} \quad \underline{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}$$

we get that $\Phi_{\underline{X}}(\underline{t}) = \mathbb{E}\left[e^{i\cdot\underline{t}'\cdot\underline{X}}\right]$ where $'$ stands for transpose. Now with $\underline{\eta} = \mathbb{E}(\underline{X})$ we get

$$\underline{t}' \cdot \underline{\eta} = \left[\ t_1, \cdots, t_n\ \right] \cdot \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} = \sum_{i=1}^{n} t_i \eta_i, \quad \text{which is a scalar.}$$

Hence generalizing the above example, the characteristic function of $\underline{X}$, a Gaussian random vector, is

$$\Phi_{\underline{X}}(\underline{t}) = \mathbb{E}\left[e^{i\underline{t}'\cdot\underline{X}}\right] = e^{\left[i\underline{\eta}'\cdot\underline{t} - \frac{1}{2}\underline{t}'\cdot\,\mathbf{A}\cdot\underline{t}\right]}$$

where $\mathbf{A}$ is the variance-covariance matrix.

Example A.2 *Let N be a Poisson random variable with parameter λ i.e.*

$$\mathbf{P}[N = n] = \frac{e^{-\lambda} \cdot \lambda^n}{n!} \qquad n = 0, 1, \cdots$$

Then the characteristic function of N is :

$$\Phi_N(t) = \sum_{n=0}^{\infty} e^{itn} \frac{e^{-\lambda} \cdot \lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{[\lambda \cdot e^{it}]^n}{n!} = e^{-\lambda} \cdot e^{\lambda e^{it}} = e^{\lambda[e^{it}-1]}, \qquad t \in \mathbb{R}.$$

Definition A.5 (Cumulant generation function) *The characteristic function $\Phi_X(s)$ of a random variable X satisfies the condition that $\Phi_X(0) = 1$ and Φ_X is continuous at $s = 0$. So $\Phi_X(s) \neq 0$ in a neighborhood of $s = 0$ and one can define a continuous version of the logarithm of Φ_X as follows: $\Psi_X(s) = \log \Phi_X(s)$ with $\Psi_X(0) = 0$.*

By taking the derivatives of cumulant generating function with respect to s one can directly obtain all the centered moments of X.
Cumulant generating function of sum of a number of independent random variables is the sum of the cumulant generating function of each random variable.

Theorem A.1 (Levy continuity theorem) *Let $\{\Phi_n; n \in \mathbb{N}\}$ be a sequence of characteristic functions and there exists a function $\Psi : \mathbb{R} \to \mathbb{C}$ such that for all $t \in \mathbb{R}$, $\Phi_n(t) \to \Psi(t)$ as $n \to \infty$ and Ψ is continuous at 0 then Ψ is the characteristic function of a probability distribution.*

Theorem A.2 (Glivenko's Theorem) *If $\{\Phi_n; n \in \mathbb{N}\}$ and Φ are characteristic functions of probability distributions $\mathcal{P}_{X_n}$ and $\mathcal{P}_X$ (respectively), where $\mathcal{P}_{X_i} = \mathbf{P} \circ X_i^{-1}$, then for all $t \in \mathbb{R}$ and for each $n \in \mathbb{N}$, $\Phi_{X_n}(t) \to \Phi_X(t) \Rightarrow \mathcal{P}_{X_n} \to \mathcal{P}_X$, weakly, as $n \to \infty$.*

A.2 Notions of Convolution

Given random variable X and Y and a function $z = g(x, y)$ we form a new random variable

$$Z = g(X, Y).$$

We want to find the density and distribution of Z in terms of the same quantities for X and Y. For real z define $D_z = \{(x, y) \mid g(x, y) \leq z\}$ then it satisfies that

$$\{Z \leq z\} = \{g(X, Y) \leq z\} = \{(X, Y) \in D_z\},$$

so that

$$F_Z(z) = \mathbf{P}[Z \leq z] = \mathbf{P}[(X, Y) \in D_z] = \iint_{D_z} f_{XY}(x, y)\text{dxdy}.$$

Thus to find F_Z it suffices to find the region D_z, for every z, and then evaluate the above integral. For instance, let

$$Z = g(X, Y) = X + Y,$$

then

$$F_Z(z) = \iint_{x+y\leq z} f_{XY}(x, y)\text{dxdy}.$$

So

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_{XY}(x, y)\text{dxdy}.$$

Using Leibnitz rule for differentiating an integral, we get

$$\begin{aligned} f_Z(z) = \frac{d}{dz}[F_Z(z)] &= \frac{d}{dz}\left[\int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_{XY}(x, y)\text{dxdy}\right] \\ &= \int_{-\infty}^{\infty} f_{XY}(z - y, y)\text{dy}, \end{aligned}$$

and with the independence of X and Y this implies

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y)\text{dy} \tag{A.4}$$

which is known as the convolution of f_X and f_Y.

Before giving the formal definition of convolution and stating some results, it is necessary to state *Fubini's theorem.*

Theorem A.3 (Fubini's theorem) *Let* $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ *and* $(\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ *be probability spaces and consider the product space* $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mathbf{P}_1 \times \mathbf{P}_2)$ *where* $\mathbf{p}_1 \times \mathbf{P}_2$ *is the product* *measure. Suppose that (X_1, X_2) is a two–dimensional random variable and g is a $\mathcal{F}_1 \times \mathcal{F}_2$ measurable function such that g is non–negative or integrable. Then*

$$\begin{aligned}
\mathbb{E}\left[g(X,Y)\right] &= \iint_{\Omega_1 \times \Omega_2} g(X_1, X_2)\mathrm{d}\left(\mathbf{P_1} \times \mathbf{P_2}\right) \\
&= \int_{\Omega_1} \left(\int_{\Omega_2} g(X_1, X_2)\mathrm{d}\mathbf{P_2} \right) \mathrm{d}\mathbf{P_1} \\
&= \int_{\Omega_2} \left(\int_{\Omega_1} g(X_1, X_2)\mathrm{d}\mathbf{P_1} \right) \mathrm{d}\mathbf{P_2} \\
&= \iint_{\mathbb{R}^2} g(x_1, x_2) F_{X_1 X_2}(\mathrm{dx_1 dx_2}) \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x_1 x_2)\mathrm{dF_{X_1}} \right) \mathrm{dF_{X_2}} \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x_1 x_2)\mathrm{dF_{X_2}} \right) \mathrm{dF_{X_1}}
\end{aligned}$$

In the settings of *Fubini's Theorem* let us consider F_1 and F_2 as the marginal distributions of X_1 and X_2 respectively.

Definition A.6 (Convolution) *The convolution of the distribution functions F_1 and F_2 is defined as*

$$F_1 * F_2(x) = \int_{-\infty}^{\infty} F_1(x-y)\mathrm{dF_2}(y), \qquad \forall x \in \mathbb{R}$$

Remark A.2 The above convolution describes the distribution of $X_1 + X_2$ as shown by the next theorem.

Theorem A.4

$$F_{X_1+X_2}(u) = F_1 * F_2(x) = \int_{-\infty}^{\infty} F_1(u-y)\mathrm{dF_2}(y)$$

Proof. We need to apply the *Fubini's theorem*

$$\begin{aligned}
F_{X_1+X_2}(u) &= \mathbf{P}\left(X_1 + X_2 \leq u\right), \qquad u \in \mathbb{R} \\
&= \iint_{x+y \leq u} \mathrm{d}\left(\mathrm{F_1} \times \mathrm{F_2}\right)(x,y) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{u-y} \mathrm{d}\left(\mathrm{F_1} \times \mathrm{F_2}\right)(x,y) \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{u-y} \mathrm{dF_1}(x) \right) \mathrm{dF_2}(y) \\
&= \int_{-\infty}^{\infty} \mathrm{F_1}(u-y)\mathrm{dF_2}(y).
\end{aligned}$$

□

Corollary A.1 *If X_2 is absolutely continuous with density f_2 then*

$$F_{X_1+X_2}(u) = \int_{-\infty}^{\infty} \mathrm{F}_1(u-y) f_2(y) \mathrm{dy}, \qquad u \in \mathbb{R}.$$

If X_1 is also absolutely continuous with density f_1 then the density of the sum can be expressed as

$$f_{X_1+X_2}(u) = \int_{-\infty}^{\infty} f_1(u-y) f_2(y) \mathrm{dy}, \qquad u \in \mathbb{R} \tag{A.5}$$

Definition A.7 (Bounded and unbounded variation) *The total variation of a right continuous function with left limit is defined as:*

$$\| f \|_{TV} := \sup \left\{ \sum_{j=1}^{j=n} | f(t_j) - f(t_{j-1}) | ; 0 = t_0 < t_1 < \cdots < t_n = t, \ n \in \mathbb{N} \right\}.$$

Clearly for an increasing function on $[0, t]$ with $f(0) = 0$ this is just $f(t)$ and for a difference $f = g - h$ of two increasing functions with $f(0) = g(0) = 0$ the total variation is at most $g(t) + h(t) < \infty$. Such functions are known as functions of finite or bounded variation. In case when the total variation of a function is infinite, the function is known as "of unbounded variation".

A.3 Jumps Caglad and Cadlag Functions

A function

$$f : [0, T] \to \mathbb{R}$$

is said to be *cadlag* if it is right continuous with left limits (*rcll*), i.e. for each $t \in [0, T]$ the limits

$$f(t^-) = \lim_{\substack{s \to t, \\ s < t}} f(s), \qquad f(t^+) = \lim_{\substack{s \to t, \\ s > t}} f(s)$$

exist and

$$f(t) = f(t^+).$$

So clearly any continuous function is cadlag but cadlag functions can have discontinuities. If t is a point of discontinuity of f, then the *jump* of f at t is defined and denoted as

$$\triangle f(t) = f(t) - f(t^-).$$

Cadlag functions can not jump around too wildly. A cadlag function can have at most a countable number of discontinuities, i.e.

$$\{t \in [0,T] : f(t) \neq f(t^-)\}$$

is finite or at most countable.

Also for any $\epsilon > 0$, the number of discontinuities(jumps) on $[0,T]$ should be finite.

So, after all, a cadlag function on $[0,T]$ has a finite number of large jumps (larger than ϵ) and possibly an infinite but countable number of small jumps.

In the definition the function is right continuous at jump times because we have defined it so. If it were defined in the other way, i.e. at the jump times t_i, the value of $f(t)$ would be defined as the *value before the jump* $\big(f(t_i) = f(t_i^-)\big)$ then the function would have been known as *Caglad*. Clearly this is not that suitable for modelling uncertainty.

Example A.3 *Consider a step function having a jump at some point T_0 whose value at T_i is defined to be the value after the jump i.e.*

$$f(T_0) := \mathbb{I}_{[T_0,T)}(t). \quad \textit{Hence} \quad f(T_0^-) = 0 \quad \textit{and} \quad f(T_0^+) = f(T_0) = 1$$

which implies that

$$\triangle f(T_0) = f(T_0) - f(T_0^-) = 1.$$

In general given a continuous function $g : [0,T] \to \mathbb{R}$ and constants $f_i, i = 0,1,2\cdots(n-1)$ with $t_0 = 0 < t_1 < t_2 < \cdots < t_n = T$, the following function is cadlag;

$$f(t) = g(t) + \sum_{i=0}^{n-1} f_i \mathbb{I}_{[t_i,t_{i-1})}(t) \qquad t \in \mathbb{R}.$$

In the above example, g is the continuous component of f to which jumps have been added. Jumps of f occur at t_i with $\triangle f(t_i) = f_i - f_{i-1}$. This is just a typical example . In practice not every Cadlag function has such a neat decomposition into continuous and jump parts.

Cadlag functions are, therefore, natural models for the sample paths of jump processes.

A.4 Stochastic Process

A random variable X is a mapping between the sample space Ω and the real line $\mathbb{R}$, i.e.

$$X : \Omega \to \mathbb{R}.$$

A stochastic process is a mapping from the sample space into an ensemble (collection) of time functions. To every $\omega \in \Omega$ there corresponds a function of time (a sample function) $X(t,\omega)$. It is usual to drop the sample space variable ω and to write simply $X(t)$.

For a fixed $t = t_0$ the quantity $X(t_0,\omega)$ is a random variable mapping Ω into the real line and for fixed ω_0 the quantity $X(t,\omega_0)$ is a well defined non-random function of time. Also for fixed t_0 and ω_0, $X(t_0,\omega_0)$ is just a real number.

Example A.4 *Consider the coin-toss experiment. So we have*

$$\Omega = \{H, T\}.$$

Then if we define $X(t,H) = \sin(t)$ *and* $X(t,T) = \cos(t)$ *we get a stochastic process. Here although the sample paths are continuous the process itself is discrete because* ω *takes discrete values.*